

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 12:17:11 ; Search time 10785 Seconds
(without alignments)

10697.393 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agtctctgcgcagaaag.....atctgaagctcactattcag 2381

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hgt:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736.8	30.9	2847	5	AY254202 Gallus ga
2	85	3.6	923	5	BX932340 Gallus ga
3	81	3.4	899	5	BX930518 Gallus ga
4	78.8	3.3	896	5	BX933597 Gallus ga
5	64.2	2.7	1141	6	AX083744 Sequence
6	62	2.6	3013	10	AY254202 Rat intesti
7	62	2.6	230897	2	AC130496 Rattus no
8	57.8	2.4	1141	6	AX083744 Sequence
9	56.6	2.4	5039	10	MUSFABPI
10	51.8	2.2	234081	3	PFMAL1301
11	50	2.1	564	10	RATFABPX
12	49.8	2.1	619	5	XELIFABP
13	49.8	2.1	5204	6	AX083744 Sequence
14	49.8	2.1	5204	9	HUMFABP
15	49.8	2.1	200000	2	AC008077 Homo sapi
16	49.2	2.1	2000	6	AX655393 Sequence
17	49	2.1	136061	2	AC117821 Mus muscu
18	49	2.1	202804	10	AC021630 Mus muscu
19	49	2.1	249943	3	AE014823 Plasmodiu

20	48.4	2.0	564	10	RATFABPI
C 21	48.4	2.0	832	6	AR415176 Sequence
C 22	48.4	2.0	832	6	AX972010 Sequence
C 23	48.4	2.0	832	6	BD110729 EST and e
C 24	48.4	2.0	18876	3	CEF32G8
C 25	48.4	2.0	178045	2	CR762472
C 26	48.4	2.0	197110	9	AC104306
C 27	48.2	2.0	132254	3	AC116330 Dictyoste
C 28	48.2	2.0	132254	3	AC116330 Dictyoste
C 29	48	2.0	191840	9	AC092656 Homo sapi
C 30	47.8	2.0	110000	3	AC116984_1
C 31	47.8	2.0	137174	2	BX276186 Danio rer
C 32	47.8	2.0	162810	2	CR392028 Danio rer
C 33	47.6	2.0	5302	3	AC116961 Dictyoste
C 34	47.6	2.0	143364	2	BX324220 Danio rer
C 35	47.6	2.0	174176	5	AL772340 Zebrafish
C 36	47.6	2.0	177623	2	CR394571 Danio rer
C 37	47.4	2.0	110000	2	PFMAL13_09
C 38	47.4	2.0	318221	2	PFMAL13P3
C 39	47.2	2.0	165797	2	CR847851 Danio rer
C 40	47	2.0	2636	6	CO593303 Sequence
C 41	47	2.0	81181	2	AC019668 Drosophil
C 42	47	2.0	176969	3	AC010707 Drosophil
C 43	47	2.0	181815	5	BX248120 Zebrafish
C 44	47	2.0	184977	5	BX649334 Zebrafish
C 45	47	2.0	192540	3	AC010846 Drosophil

ALIGNMENTS

RESULT 1	AY254202	2847 bp	DNA	linear	VRT 22-APR-2003
LOCUS	Gallus gallus intestinal fatty acid-binding protein gene, complete cds.				
DEFINITION	AY254202				
ACCESSION	AY254202				
VERSION	AY254202.1				
KEYWORDS	Gallus gallus (chicken)				
SOURCE	Gallus gallus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
REFERENCE	1 (bases 1 to 2847)				
AUTHORS	Wang, Q., Li, H., Wang, Y. and Zhao, J.				
TITLE	Cloning and characterization of chicken I-FABP gene				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 2847)				
AUTHORS	Wang, Q., Li, H., Wang, Y. and Zhao, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-MAR-2003) Animal Science & Technology, Northeast Agricultural University, 59 Mucai Street, Harbin, Heilongjiang 150030, China				

FEATURES	Location/Qualifiers
source	1..2847
	/organism="Gallus gallus"
	/mol_type="genomic DNA"
	/db_xref="taxon:9031"
	/tissue_type="blood"
mRNA	/note="Breed: Arber Acres broiler"
	join(<21..87,720..892,1465..1572,2416..>2466)
CDS	/product="intestinal fatty acid-binding protein"
	join(21..87,720..892,1465..1572,2416..>2466)
	/note="FABP"
	/codon_start=1
	/product="intestinal fatty acid-binding protein"
	/protein_id="AAP13101.1"
	/db_xref="GI:30060212"
	/translation="MAPNGTWIKENYEKFEWAMGVNMRKLGALHNKLTIQDD GNKFLVKSSNFRITDIEFTLGVSFEYSLADGTLSGNWLNLEGNKLVGFTTRKNGKV LTAYPEIVGSELITQTVYVGEAKRIFKKE"
ORIGIN	


```
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST640b17"
/clone_lib="CSEQCHN56"
/dev_stage="adult"

ORIGIN
Query Match 3.4%; Score 81; DB 5; Length 899;
Best Local Similarity 100.0%; Pred. No. 9.4e-07;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1614 TACAGACAGAAAGATGGCATTTACGGTACTTGGGAAATAGAGAAAAATGAGAACTATGA 1673
Db 11 TACAGACAGAAAGATGGCATTTACGGTACTTGGGAAATAGAGAAAAATGAGAACTATGA 70

Qy 1674 AAAATTTCATGGAAGCAATGGG 1694
Db 71 AAAATTTCATGGAAGCAATGGG 91

RESULT 4
BX933597 896 bp mRNA linear VRT 02-FEB-2004
LOCUS Gallus gallus finished cDNA, clone CHEST153f19.
ACCESSION BX933597
VERSION BX933597.1 GI:41634125
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 896)
AUTHORS Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
DIRECT SUBMISSION
TITLE Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
JOURNAL CB10 ISA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from small intestine, and poly A-trimmed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
FEATURES
source Location/Qualifiers
1..896
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST153f19"
/clone_lib="CSEQCHL18"
/dev_stage="adult"

ORIGIN
Query Match 3.3%; Score 78.8; DB 5; Length 896;
Best Local Similarity 97.6%; Pred. No. 2.8e-06;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1613 CTACAGACAGAAAGATGGCATTTACGGTACTTGGGAAATAGAGAAAAATGAGAACTATG 1672
Db 7 CGAGACAGAGAAAGATGGCATTTACGGTACTTGGGAAATAGAGAAAAATGAGAACTATG 65

Qy 1673 AAAAATTCATGGAAGCAATGGG 1694
Db 67 AAAAATTCATGGAAGCAATGGG 88
```

```
RESULT 5
AX083744/c
LOCUS AX083744 1141 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 22 from Patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kunst,L. and Clemens,S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: WO 0111061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES
source Location/Qualifiers
1..1141
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
/promoter
/notes="consensus sequence of A.t., L.a., and B.n. FAE1 promoters"
ORIGIN
Query Match 2.7%; Score 64.2; DB 6; Length 1141;
Best Local Similarity 12.7%; Pred. No. 0.0034;
Matches 136; Conservative 379; Mismatches 550; Indels 10; Gaps 3;

Qy 830 TTTTCTGCATTATCCTTCAACATTTAAACCTGGGATCTATGGAATCAACACGTTGGG 889
Db 1129 TTKTKYKANNNNNNNGKDNEMDATKSGATGTAWTNKAGATWCWYWTGTN 1070

Qy 890 TAAATTCATCTAGCAGCATCACTACTGTAGGAATGGACAGAAACAGACATTCAC 949
Db 1069 RRCMRTYAMRTWYTRSNANWSCATKBMMWMTWKYATKYRTAWYAMCAWRNNNNMCATN 1010

Qy 950 GAATGGCTATATATAGAGATACGTAGAGAGTCTCCTGAATTTAGACTACCTATTAAA 1009
Db 1009 GYAKSCATNNAMMYATTRWAAAYAAAKWARWAGNNRMRYGAAAGNKWGMCAATMBGWWA 950

Qy 1010 GAGTGAGGACACGAATGGAGATATCATCGCAATTTCTGTAGCTCAGACATAGCGAA 1069
Db 949 DTAGKWCNNNNNNWTTDVRMAMKAKNNNNNNAYTACYNRAATNNKGMATWKKWTHGAH 890

Qy 1070 GGTTCGTGAAATGGAACCGAG-----TTTCCAAACTACCTGTGGATGTTTCAGTGGAT 1122
Db 889 SKRTRRHTRTCRRTKYNNNNNNNARTVYVYTHAARRWMAWMTRTNNNNNNNNNACRNT 830

Qy 1123 CCTTCATCTCATGCTTATTATGTGGAGTAGAATAGATTCTCACCATAATTAGATGACAA 1182
Db 829 RTWABWKEHSWNNNNNNNNNNNNNTWCHYTTANABBCYRANNNNAARMAARTCNMYH 770

Qy 1183 AGCAGAGATTGTGGTTTATCTGTCGGTAAATACGCTTTTCTCCAGTTGTATAAGACCC 1242
Db 769 AATTTTHTDWCYKWTWNTWYDMWMTTBTTRMTTSTNNNNNNNNNNNNNNNNNNNNNN 710

Qy 1243 TCCACACAGTATAAAGTCTCTATGCAACAAAGAAATGTCAATACATCTCTTCTAGTCTCAT 1302
Db 709 WKAYAHATNNMGCWNNNTDARRTNNNTTVRRRWMNTKTRWYSTTTRHHYTCATNNNNN 650

Qy 1303 TATTATTTTCATTAGATACCGGTTTTTTTACTACAACCTCAATAAGATGAACAGATGAA 1362
Db 649 NNNNNNNNNNNSCCTCTRMWMTWTKGDMTTRKVKWRDRTTCTTVDVWADSWWVWYAN 590

Qy 1363 TGGGTTAGTGACGTGTTTATAAGAGAGAGTAAATAAG--ATACTATCATCATTTTGGAGCAA 1420
Db 589 WRCRDVTVTRNTYCKSYAHSYKWSNNAMWYRRYSARNWSSMAKWTTRNNNNWMBGBVR 530

Qy 1421 TAAGGGAGGAGAGATTTCAGCAACAGTGTGCTTACAAAGTGGAAAAACAAGTTAAACTAAA 1480
Db 1421 TAAGGGAGGAGAGATTTCAGCAACAGTGTGCTTACAAAGTGGAAAAACAAGTTAAACTAAA 1480
```


/note="consensus sequence of A.t., L.a., and B.n. FAE1 promoters"

ORIGIN

Query Match 2.4%; Score 57.8; DB 6; Length 1141; Best Local Similarity 10.9%; Pred. No. 0.077; Matches 82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;

Qy 1565 TGTTCATAAATCTCTTTCGAAAGCTCTGCTACTTACCAGAACTCTGCCTACAGACAGAA 1624
Db 201 DRDWSBKRMNYGBWKNWSYDVYYWVDDCKEKRWVRTRGRMRYMVAVTAHR 260
Qy 1625 AGATGGCATTTACGGTACTTGGAAATAGAGAAAATGAGAACTATGAAAATTCATGG 1684
Db 261 RRYNNGWTBAYVRWTTNNNNNAKAMCKRKYGWNRABVNSTCTTWKSKTKTKRTSC 320
Qy 1685 AAGCAATGGGTAAGCCCTCTTTTGAATGCCTCTTAAAGCAGGATACCACTACGGCG 1744
Db 321 WANNCRAGDANKDKHWKWAAGVYNNNNNNNTYKKARHARWDVWHSKWKHAN 380
Qy 1745 GAATACAAACTTAAGCTGTTTCATGAACCTACCATCTGCTAACTGCTCTTGTGTCGC 1804
Db 381 AAHSYRKKTBYKRTKTVNNNGTTWKRWWYKXMDMBCTYNNNNNGGRTYYGWTK 440
Qy 1805 TATTTGCCCTTGCACATTCGCCCTGCACCTTATTTGAAAGACTCTATAGAGGGGATAC 1864
Db 441 NKQWYTYKWKANCKRWADHKTCCTHNTTWKMTYWNVCYKSMGNTGKSHRBAAYV 500
Qy 1865 AAGGAAGAAAACATCTCGATTTTATTTGATTCGCGATAATCTTATGCAATTTAGCTAAT 1924
Db 501 TWYWWRRYAHANNNDWYKACTWYKVBVCWKWNNYAAYTKSSNNTYSRYRWKT 560
Qy 1925 CCAGTAGAGCATTCCAGCAGAAATTAATGAATATATGTAAGAAATATTTTGA 1984
Db 561 NNSWRWSDTRKSMGRANNYARABHYGYKNTRWBSHTWBHBRAGAAYWMBMYBAK 620
Qy 1985 TAAGACTGTTGAAAATTAACAAGAGGAAATTCGTGCTCCAGTTTGGCAGAACAC 2044
Db 621 HCMKAWYKAKYAGAGGNNNNNNNNNNNNNNATCARDYYAASRWYAMANAKEYYK 680
Qy 2045 ACATGATTCAGTCAATTTTAACTGCTAGTGTCTTAACTGTGTAACAACTGCGCTGT 2104
Db 681 BAANNAYTHANNWGCWNNATDTRTMKNNNNNNNAGTWKNNNNNAKNAKNAKNA 740
Qy 2105 AATATGATGTAACATACTATCTCTAGTTGGATAGTAGTGTGTAATACAGGCTGAACAC 2164
Db 741 AVKAAKHWKANKWAMRGHDAABATTDKRNGAYTKYTTNNNNNTYRGVVTNTAARD 800
Qy 2165 TGCTCAGTGAAGGTGAGAGAGTAAAGTCTGAGTCAAGATCTCGGCTAAGCTCCC 2224
Db 801 GWANNNNNNNNNNNNNNGSDMVWYWAYANYGTTNNNNNNNNNNNNAYAWTKWYTTDDR 860
Qy 2225 TCAACTACAGAAAGTCACATAAATAAGCAACATGATGTTCTATTTTGTCTCTCT 2284
Db 861 WRBAYTNNNNNNMYGAYDYAYYMSDTCDAWMKWDATKNNATTYNGTAWRTNNN 920
Qy 2285 GTTGTGATTAATGATTAATTAATTTT 2316
Db 921 NNNNTMTKYBYBHAANNNNNNNGKMCCTAHTWW 952

RESULT 9
MUSFABPI
LOCUS
DEFINITION Mouse Fabpi gene, exons 1-4.
ACCESSION M5033
VERSION M5033.1 GI:193218
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)

AUTHORS Green,R.P., Cohn,S.M., Sacchettini,J.C., Jackson,K.E. and Gordon,J.I.
TITLE The mouse intestinal fatty acid binding protein gene: nucleotide sequence, pattern of developmental and regional expression, and proposed structure of its protein product
JOURNAL DNA Cell Biol. 11 (1), 31-41 (1992)
MEDLINE 92153306
PUBMED 1739433
REFERENCE 2 (bases 1 to 5039)
AUTHORS Green,R.P., Sacchettini,J.C., Jackson,K.E., Cohn,S.M. and Gordon,J.I.
TITLE The mouse fatty acid binding protein gene (Fabpi): nucleotide sequence analysis and predictions of the structure of its protein product
JOURNAL Unpublished (1991)
COMMENT Original source text: Mus musculus (strain DBA/2J) Adult Liver DNA.
FEATURES
Location/Qualifiers
1..5039
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="DBA/2J"
/db_xref="taxon:10090"
/tissue_type="Liver"
/dev_stage="Adult"
467..547
/note="Domain I: homology with human I-FABP promoter region"; putative
524..537
/note="one of two repeats homologous to consensus repeat found in number of other lipid binding proteins which are expressed in small intestine"; putative"
727..817
/note="Domain II: homology to human I-FABP promoter region"; putative"
870..966
/note="Domain III: homology with human I-FABP promoter region"; putative"
884..897
/note="Second of two repeats homologous to consensus repeat"; putative"
937..942
/note="putative"
967..1069
/note="putative"
/label=Exon1
2326..2497
/note="putative"
/label=Exon2
4291..4398
/note="putative"
/label=Exon3
4818..5039
/note="putative"
/label=Exon4

ORIGIN

Query Match 2.4%; Score 56.6; DB 10; Length 5039;
Best Local Similarity 58.5%; Pred. No. 0.11; Matches 121; Conservative 0; Mismatches 79; Indels 7; Gaps 1;

Qy 1516 TGAGCTTTAGCCAGCCACATCATGTAAATTTGCTTCTCGATAAGCCTGTTCAATAAT 1575
Db 884 TGAACCTCGAATCTCCATCATCAGTATGAATTTGGTTTGAAGATAGAATAAGATAAT 943
Qy 1576 TCTCTTTGCAAGCTCTGCTACTTACCAGAAGTC-----TGCCTCAGACAGAAAGAT 1628
Db 944 TCTCTCTAGTGACAGGACTGGACCTCTCTCTTCTTAGAGACACACACAGCTGAGATCAT 1003
Qy 1629 GGCATTTAACGGTACTTGGAAAAATAGAGAAAATGAGAACTATGAAAAATTCATGAAGC 1688
Db 1004 GGCATTCGACGGCAGCTGGAAAGTAGACCGGAACGAGAACTATGAAAAGTTTCATGGAA 1063
Qy 1689 AATGGGTAAGCCTTACTTTTTTTGAATG 1715

Db 1 ACAGTCGACATCATGGCAATTTGATGGCACTTGGAAAGTAGACCGGAATGAGAACTATGAA 60

QY 1675 AAATTCATGGAAGCAATGGGTA 1696

Db 61 AAGTTCATGGAGAAATGGGCA 82

RESULT 12
XELIFABP
LOCUS
DEFINITION
Xenopus laevis intestinal fatty acid binding protein mRNA, complete cds.

ACCESSION
L19946

VERSION
L19946.1 GI:3111053

KEYWORDS
fatty acid; fatty acid binding protein; hormone-dependent regulation; metamorphosis; thyroid hormone

SOURCE
Xenopus laevis (African clawed frog)

ORGANISM
Xenopus laevis

REFERENCE
1 (bases 1 to 619)
Shi, Y.B. and Hayes, W.P.
Thyroid hormone-dependent regulation of the intestinal fatty acid-binding protein gene during amphibian metamorphosis
Dev. Biol. 161 (1), 48-58 (1994)

JOURNAL
MEDLINE
PUBMED
94123868

REFERENCE
2 (bases 1 to 619)
Shi, Y.B.
Direct Submission
Submitted (24-JUN-1993) Yun-Bo Shi, Laboratory of Molecular Embryology, Building 6, Rm. B1A02, NICHD/NIH, Bethesda, MD 20892, USA

COMMENT
Original source text: Xenopus laevis cDNA to mRNA.

FEATURES
source
1..619
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cell_type="epithelial cell"
/tissue_type="intestine"
/dev_stage="tadpole"
/tissue_lib="Lambda Zap"
41..439
/codon_start=1
/product="fatty acid binding protein"
/protein_id="AAC38012.1"
/db_xref="GI:3111054"
/translation="MAPDGTWKVDRSENKFMVGVNIVKELGAHDKLVIIQDD
GNFTVKESSTFRNIEIKFTLAQPFYSLDGTELGAWFLQNLQTLFTTRDNGKV
LQTRQLIGDELVTGYBETSKRIPKRG"

CDS

ORIGIN
Query Match 2.1%; Score 49.8; DB 5; Length 619;
Best Local Similarity 68.3%; Pred. No. 4.2; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 32

QY 1595 TACTACAGAGTCTCCCTACAGACAGAAGATGCGATTAAACGGTACTTGGAAATAG 1654

Db 9 TGCCTAGCCACACAGCTCTCCCAATCATCATCGGCTTTGTGGAACITGGGAAGTTG 68

QY 1655 AGAAAATGAGAACTATGAAAATTCATGAAGCAATGGGT 1695

Db 69 ACAGAGTGAACATGATGAGAAATTCATGAAGTTATGGGT 109

RESULT 13
AX771590
LOCUS
DEFINITION
Sequence 303 from Patent WO03004646.

ACCESSION
AX771590

VERSION
AX771590.1 GI:32438388

KEYWORDS
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1
O'Mahony, D.J., Byrne, D., Brayden, D., Lambkin, I. and Higgins, L.
Genetic analysis of Peyer's patches and M cells and methods and compositions targeting Peyer's patches and M cell receptors
Patent: WO 0304646-A 303 16-JAN-2003;
ELAN CORPORATION, PLC (IE)
Location/Qualifiers
source
1..5204
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.1%; Score 49.8; DB 6; Length 5204;
Best Local Similarity 59.2%; Pred. No. 3.1;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCCACATCATCTAAATTCCTTCTGATAAGCCTGTTTCATAAAT 1575

Db 971 TGAACCTTAAAGCTTCACATCATCAGATGAAAGTTGGTTCAAGATAGAAATATAATAAT 1030

QY 1576 TCTC-----TTTGCAAAGCTCTGCTACTTACCAGAGCTGCTCCCTACAGACAGAAAGA 1627

Db 1031 TCTCGCCCAAGGACAGACCTGAATCTCTAGCTGCTAGAGGCTGACT-CAACTGAATCA 1089

QY 1628 TGCATTTAAACGGTACTTGGAAATAGAGAAATAGAGAACTATGAGAAATTCATGGAAG 1687

Db 1090 TGGCGTTTGACACACTTGGAGGTAGACCGGAGTGAAGAACTATGACAGACTTCATGGAA 1149

QY 1688 CAATGGGTAAAGCCTTACTTTTGAATGCCT 1718

Db 1150 AAATGGGTAAAGACTTTATTTCTTGTGGCT 1180

RESULT 14
HUMFABP
LOCUS
DEFINITION
Human, intestinal fatty acid binding protein gene, complete cds, and an Alu repetitive element.

ACCESSION
M18079 J03465

VERSION
M18079.1 GI:182351

KEYWORDS
Alu repeat; fatty acid binding protein.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 5204)
Sweetser, D.A., Birkenmeier, E.H., Klisak, I.J., Zollman, S., Sparkes, R.S., Mohandas, T., Lusis, A.J. and Gordon, J.I.
The human and rodent intestinal fatty acid binding protein genes. A comparative analysis of their structure, expression, and linkage relationships
J. Biol. Chem. 262 (33), 16060-16071 (1987)

QY 1655 TACTACAGAGTCTCCCTACAGACAGAAGATGCGATTAAACGGTACTTGGAAATAG 1654

Db 9 TGCCTAGCCACACAGCTCTCCCAATCATCATCGGCTTTGTGGAACITGGGAAGTTG 68

QY 1655 AGAAAATGAGAACTATGAAAATTCATGAAGCAATGGGT 1695

Db 69 ACAGAGTGAACATGATGAGAAATTCATGAAGTTATGGGT 109

FEATURES
source
1..5204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="4q28-q31"
/note="PABPI mRNA (alt.) and introns"

prim_transcript
1028..54393

COMMENT
Original source text: Human DNA (library of T.Maniatis), clone lambda-HIFABP.
Draft entry and computer-readable sequence for [1] kindly provided by D.Sweetser, 19-JAN-1988.
Location/Qualifiers

Qy	1628	TGGCATTTAACGGTACTTGGAAATAGAGAAATGAGAACTATGAAAAATTCATGGAAG	1687
Db	10791	TGGCGTTTTCAGCACTTGGAAAGGTAGACCGGAGTGAACACTATGACAAGTTCATGGAAA	10732
Qy	1688	CAATGGGTAAAGCCTTACTTTTTCGAATGCCT	1718
Db	10731	AAATGGGTAAAGACTTTATTTCTTTGTGGCT	10701

Search completed: July 2, 2005, 17:52:45
Job time : 10795 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 10:20:51 ; Search time 1388.99 Seconds
(without alignments)

10147.594 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agctctctgcgcaaaag.....atctgaagctcactattcag 2381

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2381	100.0	2381	11 ADL90127	Adl90127 Chicken i
2	336	14.1	336	11 ADL90128	Adl90128 Chicken i
3	50.6	2.1	4590	5 AAH24065	Aah24065 Yeast AOD
4	49.8	2.1	5204	6 ABK83946	Abk83946 Human CDN
5	49.8	2.1	5204	12 ADJ74588	Adj74588 Human int
6	49.8	2.1	5204	12 ADO78215	Ado78215 Human fat
7	49.2	2.1	2000	8 ADA71938	Ada71938 Rice gene
8	47.2	2.0	337	8 ABX40871	Abx40871 Bovine ES
9	47	2.0	2636	4 ABL15880	Abli15880 Drosophil
10	46.8	2.0	662	10 ADD29587	Add29587 Mouse tum
11	46.6	2.0	610	4 AAL19953	Aal19953 Human bre
12	46.6	2.0	9742	6 ABL70479	Abli70479 Chemicall
13	46.2	1.9	6222	6 ABL32692	Abli32692 Human imm
14	45	1.9	10652	6 ABN80056	Abn80056 Human che
15	44.4	1.9	2000	8 ADA71938	Ada71938 Rice gene
16	44.2	1.9	515	3 AAL16469	Aal16469 Human col
17	44.2	1.9	2252	13 ADQ39082	Adq39082 Human SNP
18	43.6	1.8	14023	6 ABL34104	Abli34104 Human imm
19	43.2	1.8	7143	6 ABL32983	Abli32983 Human imm
20	42.8	1.8	51289	13 ABD33486	Abd33486 Human can

C 21	42.6	1.8	513	12 ACH79811	Ach79811 Human gen
C 22	42.6	1.8	611	3 AAL16371	Aal16371 Human col
C 23	42.6	1.8	2844	10 ABT31903	Abt31903 Human bre
C 24	42.6	1.8	5501	6 ABL32055	Abli32055 Human imm
C 25	42.4	1.8	230	2 AAX11681	Aax11681 Human bia
C 26	42.4	1.8	230	2 AAX112923	Aax112923 Human bia
C 27	42.4	1.8	452	3 AAA77858	Aaa77858 cDNA enco
C 28	42.4	1.8	452	4 AAI28596	Aai28596 Colon tum
C 29	42.4	1.8	452	8 ABZ32782	Abz32782 Human col
C 30	42.4	1.8	461	4 AAL10372	Aal10372 Human bre
C 31	42.4	1.8	682	11 ACN80571	Acn80571 Breast ca
C 32	42.4	1.8	5908	4 AAS45387	Aas45387 Chemicall
C 33	42.4	1.8	5908	6 ABK28232	Abk28232 DNA trans
C 34	42.2	1.8	453	6 ABL67809	Abli67809 Oesophagu
C 35	42.2	1.8	453	6 ABL62188	Abli62188 Colon ade
C 36	42.2	1.8	6621	6 ABL32912	Abli32912 Human imm
C 37	42	1.8	9095	6 ABQ67061	Abq67061 Human ang
C 38	41.8	1.8	10048	6 ABL70313	Abli70313 Chemicall
C 39	41.8	1.8	10048	6 AAS61251	Aas61251 Human gen
C 40	41.8	1.8	137870	10 ADG89426	Adg89426 Human mat
C 41	41.6	1.7	900	4 AAH34536	Aah34536 Human col
C 42	41.6	1.7	3639	6 ABL62356	Abli62356 Colon ade
C 43	41.6	1.7	3639	12 ADJ74884	Adj74884 Marker ge
C 44	41.6	1.7	3719	9 ADA10896	Ada10896 Human CDN
C 45	41.4	1.7	447	8 ABX50540	Abx50540 Bovine ES

ALIGNMENTS

RESULT 1

ADL90127

ID ADL90127 standard; DNA; 2381 BP.

XX ADL90127;

XX ADL90127;

DT 20-MAY-2004 (first entry)

XX Chicken intestinal fatty acid binding protein, iFABP, gene, 5' region.

XX Chicken; ds; intestinal fatty acid binding protein; iFABP;

KW gut specific promoter; transgenic.

XX Gallus gallus.

OS US2003177516-A1.

PN US2003177516-A1.

XX 18-SEP-2003.

PD 18-SEP-2003.

XX 14-MAR-2002; 2002US-00099663.

XX 14-MAR-2002; 2002US-00099663.

XX (HORS/) HORSEMAN N D.

XX (PRAT/) PRATT S L.

XX Horseman ND, Pratt SL;

XX WPI; 2003-898653/82.

XX New nucleic acid molecule comprising an isolated avian gut-specific gene

XX expression control region, useful for regulating heterologous nucleic

XX acids in transgenic avians, and for generating transgenic birds.

XX Claim 1; SEQ ID NO 1; 28pp; English.

XX The invention relates to an isolated nucleic acid comprising an isolated

XX avian gut-specific gene expression control region comprising as

XX ADL90127/Chicken intestinal fatty acid binding protein, iFABP, gene, 5'

XX region or ADL90128 (Chicken iFABP promoter) or its degenerate variant.

XX Also included are a recombinant DNA molecule comprising an isolate avian

XX gut-specific gene expression control region operably linked to a nucleic

XX acid insert encoding a polypeptide, an expression vector that integrates

CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (iFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector (or its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the iFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, iFABP, gene, 5' region.
XX

SQ Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match		100.0%;	Score 2381;	DB 11;	Length 2381;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2381;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGCTTCCTGCGCAGAAAAAGGCTGTGGGGTCTTGTTCCTTCCCTCACACAGCTTAAGCAAAATC	60		
Db	1	AGCTTCCTGCGCAGAAAAAGGCTGTGGGGTCTTGTTCCTTCCCTCACACAGCTTAAGCAAAATC	60		
Qy	61	CCCAAGTTCAAAACGTCGGCTGTAAAGAGGATGGCTCACTTCAAAATGAAGTGAATATG	120		
Db	61	CCCAAGTTCAAAACGTCGGCTGTAAAGAGGATGGCTCACTTCAAAATGAAGTGAATATG	120		
Qy	121	AAATAATCATAAACGAGCTCTGTGGCAGATCAGAGATAACCTCTGCTGGGCAAAAT	180		
Db	121	AAATAATCATAAACGAGCTCTGTGGCAGATCAGAGATAACCTCTGCTGGGCAAAAT	180		
Qy	181	CTTAAAGTGATAGGTAGAACAGGAGTGTTCGCAACTAAATGGACTAGATTGTACCACA	240		
Db	181	CTTAAAGTGATAGGTAGAACAGGAGTGTTCGCAACTAAATGGACTAGATTGTACCACA	240		
Qy	241	TTTGATCTTCTAGGAGACAAAGGGTCTGGAAAAACAAATTAATTTCTGGTGCAAGTCAGT	300		
Db	241	TTTGATCTTCTAGGAGACAAAGGGTCTGGAAAAACAAATTAATTTCTGGTGCAAGTCAGT	300		
Qy	301	AGCAGCTGTTTTGGGTGCAACTACAGCACTTTGTTGCAACAATAAATCTAAGTTG	360		
Db	301	AGCAGCTGTTTTGGGTGCAACTACAGCACTTTGTTGCAACAATAAATCTAAGTTG	360		
Qy	361	TTTTCTTTCTCTTTCTTAACTCTGTACAGTCTAAAGGTGAAGAGTAGCTATTGAGT	420		
Db	361	TTTTCTTTCTCTTTCTTAACTCTGTGTACAGTCTAAAGGTGAAGAGTAGCTATTGAGT	420		
Qy	421	TACTTCCTCTGCATCTCTTAGCCAGATTAGCAATTTGATTTCAAAATGAACCTGAGTGA	480		
Db	421	TACTTCCTCTGCATCTCTTAGCCAGATTAGCAATTTGATTTCAAAATGAACCTGAGTGA	480		
Qy	481	ATGGAAGCCACACTATTTTGGTATCACCAGCAAGTCTTAAATTTATAGTTATACCTTCA	540		
Db	481	ATGGAAGCCACACTATTTTGGTATCACCAGCAAGTCTTAAATTTATAGTTATACCTTCA	540		
Qy	541	GTAACACCTTTTCTGAGCTCTGGAGAAAAAGAGATTATGATAACACAGACTAGTA	600		
Db	541	GTAACACCTTTTCTGAGCTCTGGAGAAAAAGAGATTATGATAACACAGACTAGTA	600		
Qy	601	AAATTCATTAGTAGCCAAACCTGTATTCTGTGTGATGAAGCAACATTCATTTCAGCAT	660		
Db	601	AAATTCATTAGTAGCCAAACCTGTATTCTGTGTGATGAAGCAACATTCATTTCAGCAT	660		
Qy	661	TCAGGATTTACATTTTGAAGCTAATAGACAGAGATTGTCGCCGTCCATAGGAACAGA	720		
Db	661	TCAGGATTTACATTTTGAAGCTAATAGACAGAGATTGTCGCCGTCCATAGGAACAGA	720		
Qy	721	CTAACTATAATCTCGAGTTTGTAGTACAGCAGATTTAGCACAGCAAAATTTGCTCAGTTTC	780		
Db	721	CTAACTATAATCTCGAGTTTGTAGTACAGCAGATTTAGCACAGCAAAATTTGCTCAGTTTC	780		

Qy	781	AAAGTAGCACTATCTTGTGGGGAAGAGAGCTGAGCCAGTGTGCTCATTTTCTGCATT	840
Db	781	AAAGTAGCACTATCTTGTGGGGAAGAGAGCTGAGCCAGTGTGCTCATTTTCTGCATT	840
Qy	841	ATCCTTTCAACATTTAAACCTCGGATCTATGGAATCAAAACACGTTGGGTAAATTCAC	900
Db	841	ATCCTTTCAACATTTAAACCTCGGATCTATGGAATCAAAACACGTTGGGTAAATTCAC	900
Qy	901	TAGCAGCAGCATCACTACTGTAGGAATGGACAGAAACAGAGCATTCATGAAATGGCTAT	960
Db	901	TAGCAGCAGCATCACTACTGTAGGAATGGACAGAAACAGAGCATTCATGAAATGGCTAT	960
Qy	961	AATATAGAGAAATACGTAGAGGTGCTGAAATTTAGACTACCTATTAAGAGGTGAGGACA	1020
Db	961	AATATAGAGAAATACGTAGAGGTGCTGAAATTTAGACTACCTATTAAGAGGTGAGGACA	1020
Qy	1021	CGAATGGAGAAATATCATCGCAATTTCTGTAGCTCAGACACTAGACTCGAAGGTTTCTGAAA	1080
Db	1021	CGAATGGAGAAATATCATCGCAATTTCTGTAGCTCAGACACTAGACTCGAAGGTTTCTGAAA	1080
Qy	1081	CTGAACCGAGTTTCCCAAACTACTGTGGATGTTCAAGTGGATCCTCATCTCATGCTTAT	1140
Db	1081	CTGAACCGAGTTTCCCAAACTACTGTGGATGTTCAAGTGGATCCTCATCTCATGCTTAT	1140
Qy	1141	TATGTGGAGTAGAATAGATTCTCCCAATTTAGATGGACAAAGCAGAGATTTGTGTTTT	1200
Db	1141	TATGTGGAGTAGAATAGATTCTCCCAATTTAGATGGACAAAGCAGAGATTTGTGTTTT	1200
Qy	1201	ATCTGTGGGTAAATACGTTTTCTCCAGTTGTATAAAGACCCCTCCACACAGATATAAGTC	1260
Db	1201	ATCTGTGGGTAAATACGTTTTCTCCAGTTGTATAAAGACCCCTCCACACAGATATAAGTC	1260
Qy	1261	CTATGCAACAAAGAAAATGTCAATACATTTCTTTAGTCTCATTTATTTTCAATAGATA	1320
Db	1261	CTATGCAACAAAGAAAATGTCAATACATTTCTTTAGTCTCATTTATTTTCAATAGATA	1320
Qy	1321	GCCTGTTTTTTTACTACACTCAATAGATGAACAGATGAATGGTGTAGTCTGCTTTTA	1380
Db	1321	GCCTGTTTTTTTACTACACTCAATAGATGAACAGATGAATGGTGTAGTCTGCTTTTA	1380
Qy	1381	TAAAGAGAGTAATAAGATACTATCATCATTTGAGGCAATAAGGGAGGAGAGATTCAG	1440
Db	1381	TAAAGAGAGTAATAAGATACTATCATCATTTGAGGCAATAAGGGAGGAGAGATTCAG	1440
Qy	1441	CAAAACAGTGTCTTACAAGTGGAACAAAGTTAACTAAAGTGACCCCTCTCTTGACAA	1500
Db	1441	CAAAACAGTGTCTTACAAGTGGAACAAAGTTAACTAAAGTGACCCCTCTCTTGACAA	1500
Qy	1501	GATCAATGCCACAGTTGAGCTTTTAGCCAGCAGCATCATGTAAATTCCTTCTTGATA	1560
Db	1501	GATCAATGCCACAGTTGAGCTTTTAGCCAGCAGCATCATGTAAATTCCTTCTTGATA	1560
Qy	1561	AGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTACCAGAAAGTCTGCTCAGAC	1620
Db	1561	AGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTACCAGAAAGTCTGCTCAGAC	1620
Qy	1621	AGAAAGATGGCAATTTAAACCGGTACTTGGAAAAATAGAAAAATGAAGAACTATGAAAATTC	1680
Db	1621	AGAAAGATGGCAATTTAAACCGGTACTTGGAAAAATAGAAAAATGAAGAACTATGAAAATTC	1680
Qy	1681	ATGGAGCAATGGGTAAAGCTTTACTTTTTTGAATGCTTCTTAAAGCAGGATACCACTAC	1740
Db	1681	ATGGAGCAATGGGTAAAGCTTTACTTTTTTGAATGCTTCTTAAAGCAGGATACCACTAC	1740
Qy	1741	GGCGGAATCAAACTTTAAGCTGTTTCAATGAACCTACCTGCTGCTGCTGCTGCTGCTG	1800
Db	1741	GGCGGAATCAAACTTTAAGCTGTTTCAATGAACCTACCTGCTGCTGCTGCTGCTGCTG	1800
Qy	1801	CTGCTATTTTGGCTTTGCAATTTGCCCTGCACTTTATTTTGAAGAAAGCTCTATAGAGGGA	1860
Db	1801	CTGCTATTTTGGCTTTGCAATTTGCCCTGCACTTTATTTTGAAGAAAGCTCTATAGAGGGA	1860

```
Qy 1861 ATACAGGAGAGAAAAACATTCGATTTATTTGCAATTCGCGATAATCTATGCAATTTAGCT 1920
Db 1861 ATACAAGGAGAGAAAAACATTCGATTTATTTGCAATTCGCGATAATCTATGCAATTTAGCT 1920
Qy 1921 AATTCCAGTAGAGCATTCCAGCAGAAATTTAAATAGAAATATATATGTAAGGAATATTATT 1980
Db 1921 AATTCCAGTAGAGCATTCCAGCAGAAATTTAAATAGAAATATATATGTAAGGAATATTATT 1980
Qy 1981 TTGATAAGACTGTTTGAATAATACACAGAGGGAATTTGCTGCTCCAGTTTTCGAGA 2040
Db 1981 TTGATAAGACTGTTTGAATAATACACAGAGGGAATTTGCTGCTCCAGTTTTCGAGA 2040
Qy 2041 ACACACATGATTGAGTCATTTAAATGATGAGTCTTAAAGCTTTGACAACTGC 2100
Db 2041 ACACACATGATTGAGTCATTTAAATGATGAGTCTTAAAGCTTTGACAACTGC 2100
Qy 2101 CTGTAATATGATCTAAACATACTATCCTAGTTGGATAGTATGTTTCTATACAGGCTGA 2160
Db 2101 CTGTAATATGATCTAAACATACTATCCTAGTTGGATAGTATGTTTCTATACAGGCTGA 2160
Qy 2161 ACATGCTCAGTGAAGGTTGGAGAGAGTAAGACTCTGAGTCAGAAATTTCTGGGCTAAGC 2220
Db 2161 ACATGCTCAGTGAAGGTTGGAGAGAGTAAGACTCTGAGTCAGAAATTTCTGGGCTAAGC 2220
Qy 2221 TCCCTCAACTACAGAAAAAGTCAATATAAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 2280
Db 2221 TCCCTCAACTACAGAAAAAGTCAATATAAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 2280
Qy 2281 CTCTGCTTGATGTTAATGATATATATATATATATATATATATATATATATATATATATAT 2340
Db 2281 CTCTGCTTGATGTTAATGATATATATATATATATATATATATATATATATATATATATAT 2340
Qy 2341 GAAAGTTAGGAGCCACGATAATCTGAAGCTCACTATTTCAG 2381
Db 2341 GAAAGTTAGGAGCCACGATAATCTGAAGCTCACTATTTCAG 2381

RESULT 2
ID ADL90128 standard; DNA; 336 BP.
AC ADL90128;
XX
XX
XX 20-MAY-2004 (first entry)
XX
XX Chicken intestinal fatty acid binding protein, iFABP, gene, promoter.
XX
XX Chicken; ds; intestinal fatty acid binding protein; iFABP;
XX gut specific promoter; transgenic; promoter.
XX
XX Gallus gallus.
XX
XX US2003177516-A1.
XX
XX 18-SEP-2003.
XX
XX 14-MAR-2002; 2002US-00099663.
XX
XX 14-MAR-2002; 2002US-00099663.
XX
XX (HORS//) HORSEMAN N D.
XX (PRAT//) PRATT S L.
XX
XX Horseman ND, Pratt SL;
XX
XX WPI; 2003-898653/82.
XX
XX New nucleic acid molecule comprising an isolated avian gut-specific gene
XX expression control region, useful for regulating heterologous nucleic
XX acids in transgenic avians, and for generating transgenic birds.
XX
XX Claim 1; SEQ ID NO 2; 28pp; English.
XX
```

```
CC The invention relates to an isolated nucleic acid comprising an isolated
CC avian gut-specific gene expression control region appearing as
CC ADL90127(Chicken intestinal fatty acid binding protein, iFABP, gene, 5'
CC region or ADL90128 (chicken iFABP promoter) or its degenerate variant
CC Also included are a recombinant DNA molecule comprising an isolate avian
CC gut-specific gene expression control region operably linked to a nucleic
CC acid insert encoding a polypeptide, an expression vector that integrates
CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (iFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector for its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the iFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, iFABP, gene, promoter.
XX
SQ Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

Query Match 14.1%; Score 336; DB 11; Length 336;
Best Local Similarity 100.0%; Pred. No. 7.8e-70;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 ATTATTATTTCATTAGATAGCCGGTTTTTACTACAACTCAATAAGATGAACAGAAATG 1360
Db 1 ATTATTATTTCATTAGATAGCCGGTTTTTACTACAACTCAATAAGATGAACAGAAATG 60

Qy 1361 AATGGCTTAGTACTCTTTTATAAAGAGATTAAGATACATCATCTATCTTGGGCAA 1420
Db 61 AATGGCTTAGTACTCTTTTATAAAGAGATTAAGATACATCATCTATCTTGGGCAA 120

Qy 1421 TAAGGGAGGAGAGATTCAGCAACAGTGTGCTTACAAGTGAACAACTTAACTAAA 1480
Db 121 TAAGGGAGGAGAGATTCAGCAACAGTGTGCTTACAAGTGAACAACTTAACTAAA 180

Qy 1481 GTGACCCCTCTCTTCAAGATCAATGCGACAGTTGAGCTTTAGCCAGCCACATATCA 1540
Db 181 GTGACCCCTCTCTTCAAGATCAATGCGACAGTTGAGCTTTAGCCAGCCACATATCA 240

Qy 1541 TGTAAATTCCTTCCTGATAAGCCTCTTCATAAATCTCTTTCGAAGCTCTGCTACTTA 1600
Db 241 TGTAAATTCCTTCCTGATAAGCCTCTTCATAAATCTCTTTCGAAGCTCTGCTACTTA 300

Qy 1601 CCAGAAAGTCTGCTCAGACAGAAAGATGGCATTTA 1636
Db 301 CCAGAAAGTCTGCTCAGACAGAAAGATGGCATTTA 336

RESULT 3
AAH24065
ID AAH24065 standard; DNA; 4590 BP.
XX
XX AAH24065;
XX
XX 29-AUG-2001 (first entry)
XX
XX Yeast AOD9604-associated DNA sequence, SEQ ID NO:1.
XX
XX Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;
XX modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;
XX lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;
XX functional food; transgenic yeast; fat/lean ratio; food use; ds.
XX
XX Saccharomyces cerevisiae.
XX
XX Key Location/Qualifiers
XX misc_feature 10
XX
```


PT Analysis of specific gene polymorphisms in clinical nucleic acid sample
PT for prediction of risk of restenosis after balloon or stent coronary
XX angioplasty.
XX
PS Disclosure; SEQ ID NO 12; 164pp; Japanese.
XX
CC The invention relates to a novel method for predicting the risk of
CC restenosis after coronary angioplasty comprising analysing specific gene
CC polymorphisms in a clinical nucleic acid sample. The method is useful for
CC the diagnosis of the genetic risk of restenosis following balloon or
CC stent coronary angioplasty. The method has high accuracy and high
CC estimation ratio. The present sequence is used in the exemplification of
CC the invention.
XX
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
Query Match 2.1%; Score 49.8; DB 12; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.26;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
QY 1516 TGAGCTTTAGCCAGCCACATCATCTGTAATTTCTCTGATAAGCCTGTTTCATAAAT 1575
Db 971 TGAACCTTTAAGCTTCCACATGATGAGTTGGTTCAAGATAAGAAATATAATAAAT 1030
QY 1576 TCTC-----TTTGCAGGCTCTGCTACTTACCAAGAGTCTGCCTACAGACAGAAAGA 1627
Db 1031 TCTCGCCCAAGGACACCTGAACTCTAGCTGCCCTAGAGGCTGACT-CAACTGAAATCA 1089
QY 1628 TGGCATTAAACGGTACTTGGAAAATAGAGAAAATGAGAACTATGAAAAATTCATGGAAG 1687
Db 1090 TGGCGTTTGACACACTTGGNAGGTAGACCGGAGTGAAAACACTATGACAAATTCATGGAAG 1149
QY 1688 CAATGGGTAGCCTTACTTTTGAATGCGT 1718
Db 1150 AAATGGGTAAGACTTTATTCTTTGTGGCT 1180
RESULT 6
ID ADO78215
AC ADO78215 standard; DNA; 5204 BP.
XX
XX ADO78215;
DT 29-JUL-2004 (first entry)
XX
XX Human fatty acid binding protein 2 Alu repeat region DNA with SNP Seq 9.
XX
KW human; ds; single nucleotide polymorphism; SNP; hereditary risk;
KW coronary spasm; tumour necrosis factor alpha;
KW NADH/NADPH oxidase p22 phox gene; apolipoprotein E; apolipoprotein C-III;
KW stromelysin 1; interleukin 6;
KW endothelium type nitrogen monoxide synthetase;
KW fatty acid binding protein 2; cardiac sudden death;
KW variant angina pectoris; myocardial infarction.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT variation replace(2445,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
XX
XX JP2004113093-A.
XX
XX 15-APR-2004.
XX
XX 25-SEP-2002; 2002JP-00280031.
XX
XX 25-SEP-2002; 2002JP-00280031.
XX
XX (NAGO-) ZH NAGOYA SANGYO KAGAKU KENKYUSHO.
XX (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUO.
XX

DR WPI; 2004-360127/34.
XX
PT Detecting genotype in nucleic acid sample useful for diagnosing risk of
PT coronary spasm, by analyzing polymorphisms in tumor necrosis factor alpha
PT gene, NADH/NADPH oxidase p22 phox gene and apolipoprotein E gene, in
PT nucleic acid sample.
XX
PS Disclosure; SEQ ID NO 9; 59pp; Japanese.
XX
CC This invention relates to a novel method for detecting a hereditary risk
CC for a coronary spasm. Specifically, it refers to the identification of
CC two or more polymorphisms and determining the genotype of the nucleic
CC acid sample in order to calculate the risk of the patient to a coronary
CC spasm. The present invention describes a risk analysis that comprises
CC identifying two or more polymorphisms occurring at any position taken
CC from -863 of the tumour necrosis factor alpha gene, 242 of the NADH/NADPH
CC oxidase p22 phox gene, -219 of the apolipoprotein E gene, -482 of the
CC apolipoprotein C-III gene, -1171 of the stromelysin 1 gene, 4070 of the
CC apolipoprotein E gene, -634 of the interleukin 6 gene, -786 of the
CC endothelium type nitrogen monoxide synthetase gene or position 2445 of
CC the fatty acid binding protein 2 gene. Accordingly, the method enables
CC prevention of cardiac sudden death resulting from variant angina
CC pectoris, coronary spasm and myocardial infarction. This polynucleotide
CC is the human fatty acid binding protein 2 gene Alu repeat region
CC containing a G2445A SNP in the sequence given in the invention. NOTE:
CC This sequence is referred to as Seqid 9 in the sequence listing, but
CC Seqid 8 also refers to a primer given in the disclosure of the
CC specification.
XX
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
Query Match 2.1%; Score 49.8; DB 12; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.26;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
QY 1516 TGAGCTTTAGCCAGCCACATCATCTGTAATTTCTCTGATAAGCCTGTTTCATAAAT 1575
Db 971 TGAACCTTTAAGCTTCCACATGATGAGTTGGTTCAAGATAAGAAATATAATAAAT 1030
QY 1576 TCTC-----TTTGCAGGCTCTGCTACTTACCAAGAGTCTGCCTACAGACAGAAAGA 1627
Db 1031 TCTCGCCCAAGGACACCTGAACTCTAGCTGCCCTAGAGGCTGACT-CAACTGAAATCA 1089
QY 1628 TGGCATTAAACGGTACTTGGAAAATAGAGAAAATGAGAACTATGAAAAATTCATGGAAG 1687
Db 1090 TGGCGTTTGACACACTTGGNAGGTAGACCGGAGTGAAAACACTATGACAAATTCATGGAAG 1149
QY 1688 CAATGGGTAGCCTTACTTTTGAATGCGT 1718
Db 1150 AAATGGGTAAGACTTTATTCTTTGTGGCT 1180
RESULT 7
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
XX ADA71938;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5263.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX

PR 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
SQ

Query Match 2.1%; Score 49.2; DB 8; Length 2000;
Best Local Similarity 10.8%; Pred. No. 0.27;
Matches 83; Conservative 322; Mismatches 360; Indels 1; Gaps 1;
QY 839 TTATCTCTCAACATTTAAACCTGGATCTATGGAATCAACACGTTGGTGAATCA 898
DB 291 KSYISWYVYRGGRGATRYGNGYNGMYKMYKTYRGYKNGKRGWAGRMWR 350
QY 899 CTTAGCAGCAGCATCACTACTGTAGGAATGAGAACAGACGATTCAGTGGGCT 958
DB 351 SMCRWSKACYVMRWWRWMTRRRWAWSKSSRTSRKCRKCRKRYKRYGYSRMSCK 410
QY 959 ATAATATAGAGATACGTAGAGGTCTGCAATTTAGACTACCTATTAAGAGTAGGA 1018
DB 411 RARMKCRSGRAWKGRCGCMTCMKSYGMYRWSKWSKRYKMYRWRKCKCSR 470
QY 1019 CACGAATGGAGAAATATCATCGCAATTTCTGTAGCTCAGACTAGACTGGAAGTTCTGA 1078
DB 471 TTMGKTRGGMGTGRCRYKRSKCRKCRRRRGRMYRMYRMYRMYRMYRMYRMYR 530
QY 1079 AACTGAACCGAGTTTCCCAACTACCTGTGGATGTTTCAGTGGATCTCTCATCTGCTT 1138
DB 531 KYSYSAARKARCMYRGKGYWAGMMKRYKRYMYKMYKMYKRYKRYKRYKRYKRYKRY 590
QY 1139 ATTATGTG-GAGTAGATAGATTTCTACCAATTTAGATGACAAACAGAGATTTGTG 1197
DB 591 MKSARKAGAKWCKRNSAWSKMSRSRCKRCKKASKRSARVAMMGMTSGSRMSRWKS 650
QY 1198 TTATCTGTTGGTAAATAGCTTTCTCCAGTTGTTATATAAGACCTCCACAGTATAA 1257
DB 651 YTCWRKWSGMSKSTCTWYMSKTYAKYGYWRYRACMYRMYRMYRMYRMYRMYRMYR 710
QY 1258 GTCTATGCAACAAAGAAATGTCAATACATTTCTTAGTCTCATTTATTTTCATTAG 1317
DB 711 WYTSSTRMAMTGMKYSGRYTSWYKCKSKWYMSWYMSWYMSWYMSWYMSWYMSWY 770
QY 1318 ATAGCCGGTTTTTACTACAACTCAATATAGATGACAGATGAATGGGTAGTACTGT 1377
DB 771 MWYRSMKTYTCTWYMSWYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 830
QY 1378 TTATAAGAGAGTAATAAGATATCATCTATTCAGGCAATAAGGGAGGAGATT 1437
DB 831 KMRMGMTGAKTRGRARKARYVWKKATWATCATKRWMTKGKGAWAKWAKWAKWAKWAK 890

QY 1438 CAGCAAAACAGTGTCTTCAAGTGGAAAAACAAGTAAACAGTAAAGTACCCCTCTTGA 1497
DB 891 AWYVYKTRTRTRYKTCWKARWGSWAYVWMMWGSAAWMMWMMWMMWMMWMMWMMWMMW 950
QY 1498 CAAGATCAATGCCACAGTGTGAGCTTTAGCCAGCAGCATCATCATGTAATGCTTCTG 1557
DB 951 KMACGRATKYMCCAGWMMYMSYSWTRTYWRTWRMMWASSRTAKRMAMMMWMMWMMW 1010
QY 1558 ATAAGCCCTGTTCAATAATCTCTTTGCAAAAGCTCTGCTACTACCA 1603
DB 1011 ARAYWAGAGCACTACACATACACAGAAATGTTGTTTATTTTCCA 1056
RESULT 8
ABX40871/c
ID ABX40871 standard; cDNA; 337 BP.
XX
XX AC ABX40871;
XX
XX DT 20-FEB-2003 (first entry)
XX
XX DE Bovine EST associated with lactation/muscle/fat deposition #6036.
XX
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
XX OS Bos Taurus.
XX
XX PN US2002137139-A1.
XX
XX PD 26-SEP-2002.
XX
XX PF 24-SEP-2001; 2001US-00960352.
XX
XX PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
PA (NATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX
XX PS Claim 2; SEQ ID NO 6036; 245pp; English.
XX
XX CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 1512 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridization between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle

CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX SQ Sequence 337 BP; 150 A; 35 C; 40 G; 112 T; 0 U; 0 Other;
Query Match 2.0%; Score 47.2; DB 8; Length 337;
Best Local Similarity 53.9%; Pred. No. 0.44;
Matches 97; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 2185 AAGAGTAAGACTCTGAGTCAGAAATCTGGGCTAAGCTCCCTCAACTACAGAAAAGTCAC 2244
Db 289 AATAAATATATATATTTAAACAAATTAAGTTACATTTATATAAAAAAATGTGAT 230
QY 2245 AATAAATAATCAACATGATTTCTATTTTGTCTCTGCTTGATTTAATGATTAT 2304
Db 229 ATTAATATATTAATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 170
QY 2305 TATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2364
Db 169 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 110

RESULT 9
ABL15880
ID ABL15880 standard; cDNA; 2636 BP.
XX AC ABL15880;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 42122.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB1777.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX Claim 1; SEQ ID NO 42122; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins (AB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2636 BP; 849 A; 414 C; 497 G; 876 T; 0 U; 0 Other;
Query Match 2.0%; Score 47; DB 4; Length 2636;
Best Local Similarity 52.3%; Pred. No. 0.97;
Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 1821 ATTGCCCTGCACCTTATTTTGAAGAACTCTATAGAGGGGAATACAGAGGAAGAAACATT 1880
Db 540 ATTGACTTCGAAATATCTTTAAATAATACCAAGCCCTGTTCTGGTTCTTCGAATTAACAAT 599
QY 1881 CTGATTTTATTTGTCATTCGATAAATCTTATGCAATTTAGCTAATTCAGTAGAGGCATTC 1940
Db 600 ATAAATATTTTAAATGGTGAATTAATTTCTGTTCTAAATAATTCAGGTATATTATTC 659
QY 1941 AGCAGAAATTTAAATAGAAATATATGTAAGGAATATTTTGTATAAGACTGTTGAAAA 2000
Db 660 TTAGGGAATAAAGTATTATTGCTGCTAGTTTATTATGTTTAAAAAAGGTATAGACAA 719
QY 2001 ATTACACAGAGGGAATTT 2019
Db 720 ATTAGTAGATGTGTAATTT 738
RESULT 10
ADD29587
ID ADD29587 standard; mRNA; 662 BP.
XX AC ADD29587;
XX DT 15-JAN-2004 (first entry)
XX DE Mouse tumour suppressor mRNA SEQ ID NO:36.
XX KW ss; mouse; tumour suppressor; cancer; cancer; cytostatic; gene therapy.
XX OS Mus musculus.
XX PN WO2003058201-A2.
XX PD 17-JUL-2003.
XX PF 31-DEC-2002; 2002WO-US041825.
XX PR 31-DEC-2001; 2001US-0345317P.
XX PA (QUAR-) QUARK BIOTECH INC.
XX PA (CLEV-) CLEVELAND CLINIC FOUND.
XX PI Feinstein E, Gudkov AV;
XX WPI; 2003-598393/56.
XX Diagnosing cancer comprises determining the polypeptide or polynucleotide
XX levels e.g., hepatic lipase, in a sample from a subject, where a higher
XX level compared to that in a subject free of cancer is indicative of
XX cancer.
XX Disclosure; SEQ ID NO 36; 272pp; English.
XX The invention relates to a novel method for diagnosing a cancer in a
XX subject, the method comprises determining, in a sample from the subject,
XX the level of at least one polypeptide, where a higher level of the
XX polypeptide compared to the level of the polypeptide in a subject free of
XX cancer is indicative of cancer. The polypeptide is selected from any of
XX the polypeptides encoded by the polynucleotides listed in the
XX specification and polypeptides which are at least 70% homologous to the
XX polypeptides. The method of the invention has cytostatic activity, and
XX may have a use in gene therapy. The method is useful in identifying
XX markers specific for one or several types of cancer, depending on the
XX tissue origin, which may be used in numerous diagnostic and prognostic
XX applications as well as cancer type-specific targets for therapeutic
XX intervention. The compounds that modulate the activity of a tumour

Db 6886 TATGTAGATTTTAAATTAATTTGATGAAGTGTGTTATTTAATAGGTAGATTTTAA 6945
QY 2171 AGTGAAGTGGAGAGAGTAACTCTGAGTCAGAAATCTGGCTAAGCTCCCTCAACT 2230
Db 6946 GGTATATGTTGGTAAATTAATTTGTTGTTTATTTAATTAATGATATTTTATATT 7005
QY 2231 ACAGAAAAGTCAATATAAATGCAACATGATGTTCTATTTGTTTTCCTGCTTGA 2290
Db 7006 ATTATTAATTTAAGATAATGTTTATAGTGAATGTTGTTATGTTTATTTTATTTT 7065
QY 2291 TGTAAATGATTAATTAATTTTATTTTATTTTATAGCGTGAATGTGAT 2335
Db 7066 TATATGTTGTTGTAATTTTGTGTTTATTTTATTTTATTTGATAGGTT 7110

RESULT 13
ABL32692
ID ABL32692 standard; DNA; 6222 BP.
XX
AC ABL32692;
XX
XX 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 665.
DE
DE Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW db.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
PR
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX
PS Claim 1; SEQ ID NO 665; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 6222 BP; 1941 A; 55 C; 1168 G; 3058 T; 0 U; 0 Other;

Query Match 1.9%; Score 46.2; DB 6; Length 6222;
Best Local Similarity 66.7%; Pred. No. 2;
Matches 66; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1623 AAGATGGCATTTAACGGTACTTTGGAAAATAGAGAAAATGAGAACTATGAAAAATTCAT 1682

Db 4958 AATTATGGCGTTTGATAGTATTTGGAAGGTAGATCGAGTGAATATGATAAGTTTAT 5017
QY 1683 GGAAGCAATGGGTAAAGCCTTACTTTTGAATGCCTTCT 1721
Db 5018 GGAATAATGGGTAAAGATTTATTTTGTGGTTTAT 5056
RESULT 14
ABN80056
ID ABN80056 standard; DNA; 10652 BP.
XX
AC ABN80056;
XX
XX 15-JUL-2002 (first entry)
XX Human chemically modified disease associated gene SEQ ID NO 73.
DE
DE Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
XX Homo sapiens.
OS
OS Synthetic.
XX
XX WO200200927-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007536.
PP
XX 30-JUN-2000; 2000DE-01032529.
PR
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130908/17.
XX
XX Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a sequence
PT of a segment of chemically pretreated DNA of genes associated with
PT development.
XX
XX Claim 1; SEQ ID NO 73; 27pp; English.
XX
XX The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in the
CC specification such as ACCPN, ADFN, or APD1 and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to derwent by the European Patent Office
XX
SQ Sequence 10652 BP; 2132 A; 601 C; 3186 G; 4733 T; 0 U; 0 Other;
Query Match 1.9%; Score 45; DB 6; Length 10652;
Best Local Similarity 47.1%; Pred. No. 4.7;
Matches 138; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy	1835	ATTTTGAAGAC	TCTATAGAGG	GGAATACAAG	GAAGAAAACA	CATTCTGAT	TTTATTATTC	1894
Db	588	ATTGATTAAT	TATTAATTAAG	TAGAGAAATG	TTTGAATTAAT	TCTAATTAAT	TGTTGT	647
Qy	1895	ATTCCGATAAT	CTTTATGCAAT	TTCAGTAG	GAGGANTTCC	AGCAGAAATTT	TAAT	1954
Db	648	TTTTTTATATA	AAAAATTATAT	TTTATTTAT	TTCGATGAAT	GAAAAAGAAAT	TAGTT	707
Qy	1955	TAGAATTATAT	GTAAAGGAAT	TATTTTGTAT	TAGACTGTTT	GAAAAAATAC	ACAAGAGG	2014
Db	708	TATGATTTTAA	TTTAATATAT	TGTTTTTAAAAA	TATATTTTTTT	TAGTTAGTTAG	TATA	767
Qy	2015	AAATTGCTG	GTCCTCCAGT	TTTGCAGACA	CACATGAT	TTCAGTCATTTT	TAACGCTAGT	2074
Db	768	TAAATTAAT	TGAGTTTTTT	TGGTTAAGTAT	GATTTAGTTG	TGATTTTAA	GACGGGAGT	827
Qy	2075	GCTTACTTTA	AGCTGTGACA	AACTGCCTG	TAAATGGAAT	TGGATGTAA	CAATACCTAT	2127
Db	828	GGTGTGTTA	GATATTTTTTT	TTTACGTGAAAT	TAGATTAAT	GAGTTATAT	TTAT	880

RESULT 15

ADA71938/c
ID ADA71938 standard; DNA: 2000 BP.

AC ADA71938;

DT 20-NOV-2003 (first entry)

Rice gene, SEO ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene: ds.
 KW

OS Orvza sativa.

AA PN WO2003000898-A1.

03-JAN-2003.

22-JUN-2001: 2001WO-IB001105.

AA 22-JUN-2001; 2001WO-IB001105.
PR

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y; PI
Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G; PI

WPI: 2003-175290/17.

AA Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.

XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 1.9%; Score 44.4; DB 8; Length 2000;

[illegible]

Search completed: July 2, 2005, 14:27:29
Job time : 1396.99 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 13:46:26 ; Search time 432.909 Seconds

(without alignments)

8999.522 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agctctctgcgcagaaagg.....atctgaagctcactattcag 2381

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	64.2	2.7	1141	4	US-09-806-708B-22
C 2	57.8	2.4	1141	4	US-09-806-708B-22
C 3	49.8	2.1	8905	4	US-09-949-016-11761
C 4	49.8	2.1	8907	4	US-09-949-016-16261
C 5	48.4	2.0	832	4	US-09-621-976-2813
C 6	44.2	1.9	515	3	US-09-385-982-474
C 7	44.2	1.9	2252	4	US-09-949-016-4519
C 8	44.2	1.9	2273	4	US-09-949-016-19
C 9	43.8	1.8	7218	1	US-08-232-463-14
C 10	43	1.8	832	4	US-09-621-976-2813
C 11	42.6	1.8	601	4	US-09-949-016-18019
C 12	42.6	1.8	601	4	US-09-949-016-18020
C 13	42.6	1.8	601	4	US-09-949-016-161279
C 14	42.6	1.8	601	4	US-09-949-016-161280
C 15	42.6	1.8	611	3	US-09-385-982-376
C 16	42.4	1.8	452	4	US-09-401-064-138
C 17	42.2	1.8	652	4	US-09-270-767-10640
C 18	42.2	1.8	1257	4	US-09-270-767-15058
C 19	42	1.8	601	4	US-09-949-016-85025
C 20	42	1.8	601	4	US-09-949-016-85176
C 21	41.6	1.7	1273	4	US-09-270-767-14731
C 22	41.6	1.7	47363	4	US-09-949-016-13420
C 23	41.2	1.7	196	4	US-09-442-054A-42
C 24	41.2	1.7	196	4	US-09-442-054A-42
C 25	40.8	1.7	601	4	US-09-949-016-85023
C 26	40.8	1.7	601	4	US-09-949-016-85024
C 27	40.8	1.7	601	4	US-09-949-016-85026

28	40.8	1.7	601	4	US-09-949-016-85174	Sequence 85174, A
29	40.8	1.7	601	4	US-09-949-016-85175	Sequence 85175, A
30	40.8	1.7	601	4	US-09-949-016-85177	Sequence 85177, A
31	40.8	1.7	109159	4	US-09-949-016-14169	Sequence 14169, A
32	40.8	1.7	109159	4	US-09-949-016-14170	Sequence 14170, A
33	40.4	1.7	601	4	US-09-949-016-85027	Sequence 85027, A
34	40.4	1.7	601	4	US-09-949-016-85178	Sequence 85178, A
C 35	40.4	1.7	601	4	US-09-949-016-189988	Sequence 189988, A
C 36	40.4	1.7	192700	4	US-09-949-016-11820	Sequence 11820, A
C 37	40.4	1.7	192704	4	US-09-949-016-17182	Sequence 17182, A
C 38	40.4	1.7	640681	4	US-09-790-988-1	Sequence 1, Appl
C 39	40.2	1.7	1055	4	US-09-806-708B-23	Sequence 23, Appl
C 40	40.2	1.7	28862	4	US-09-949-016-16793	Sequence 16793, A
C 41	40.2	1.7	360470	4	US-09-949-016-13173	Sequence 13173, A
C 42	40	1.7	601	4	US-09-949-016-146919	Sequence 146919, A
C 43	40	1.7	601	4	US-09-949-016-146920	Sequence 146920, A
C 44	40	1.7	601	4	US-09-949-016-146921	Sequence 146921, A
C 45	40	1.7	733	4	US-09-270-767-10136	Sequence 10136, A

ALIGNMENTS

RESULT 1

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL promoters
US-09-806-708B-22

Query Match	2.7%	Score 64.2;	DB 4;	Length 1141;
Best Local Similarity	12.7%	Pred. No. 4.3e-07;		
Matches	136;	Conservative 379;	Mismatches 550;	Indels 10; Gaps 3;
QY	830	TTTTCTGCATTATCTTCAACATTTAAACCTGGGATCTCTGAAATCAACACGTTGGG	889	
Db	1129	TTTKTKYKANNNNNNNGKMDNRMDATKWSATGTMTNHAHKGATMCWYWTGTN	1070	
QY	890	TAAATTTCACTAGCAGCATCACTACTGTAGGAATGACAGAACAGCATTCCT	949	
Db	1069	RMCMFTYAMRTWYTSNANWSCATKBMWMTKWTATKYRTAYAMCAWRNNMWCATN	1010	
QY	950	GAATGGGCTATATATAGAGAATACGTAGAAGGTCTCTGAATTTAGACTTACCTATTTAA	1009	
Db	1009	GYAKSCATNNMAYATTRNAAVAAAKWAGNNMRYGAAAGKNGCMMAATMBBWA	950	
QY	1010	GAGTAGGACACGAATGGAGATATCATCGCAATTTCTGTAGCTCAGCAGTACCTGAA	1069	
Db	949	DTAGKCNNNNNNNNTTDFVRMAMKAKNNNNNNNAYTACYNRAATNNKMATHMWKTGHAH	890	
QY	1070	GGTTCTCGAACTGAACCCGAG-----TTTCCAACTACCTGTGCGATGTTTCAGTGGAT	1122	
Db	889	SKRTRHHTCTCRTKYNNNNNNNNVYVYHHAARRMAWWTTRNNNNNNNNNNACRNT	830	
QY	1123	CCTTCATCTCATGCTTATTATGTGGAGTAGAATAGATTCTCACCATAATTAGAAATGACAA	1182	

Query Match	2.1%	Score 49.8	DB 4	Length 8907
Best Local Similarity	59.2%	Pred. NO. 0.0098		
Matches 125	Conservative	% Mismatches 77	Indels 9	Gaps 2
Qy	1516	TGAGCTTTAGCCAGCCATCATCATGTAATAATTGCTTTCTCGATAGAGCTGTTTCATAAT	1575	
Db	1944	TGAACCTTTAGCTTCCACATCACAGTATGAAATTGGTTTCAGATAGAAATATATAAAT	2003	
Qy	1576	TCTC-----TTTGCAAAGCTCTGCTACTTTACCAGAAAGTCTGCCTCAGACAGAAAGA	1627	
Db	2004	TCTCGCCCAAGCAGACCTGAACTCTCTAGCTGCCTAGAGGCTGACT-CAACTGAAATCA	2062	
Qy	1528	TGGCATTTAAGCGTACTTTGGAAATAGAGAAAAATAGAGAACTATGAAAAATTCATGGAAAG	1687	
Db	2063	TGGCGTTTGA CAGACATCTGGAAGGTAGACCGAGGTGAAAACCTATCAGCAAGTTCATGGAAA	2122	

RESULT 6

```
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 474
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(515)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-474

Query Match
Best Local Similarity 1.9%; Score 44.2; DB 3; Length 515;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAAGATGGCATTAAACGGTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCAT 1682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23 AATCATGGCGTTTGACGACCTTGGAGGTAGACCGGAGTGAAGAACTATGACAAATTCAT 82

QY 1683 GGAAGCAATGGGT 1695
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 GGAAGCAATGGGT 95

RESULT 7
US-09-949-016-4519
; Sequence 4519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4519
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4519

Query Match
Best Local Similarity 1.9%; Score 44.2; DB 4; Length 2252;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAAGATGGCATTAAACGGTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCAT 1682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 AATCATGGCGTTTGACGACCTTGGAGGTAGACCGGAGTGAAGAACTATGACAAATTCAT 117

QY 1683 GGAAGCAATGGGT 1695
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GGAAGCAATGGGT 130

RESULT 8
US-09-949-016-19
; Sequence 19, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 2273
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19

Query Match
Best Local Similarity 1.9%; Score 44.2; DB 4; Length 2273;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAAGATGGCATTAAACGGTACTTGGAAAAATAGAAAAATGAGAACTATGAAAAATTCAT 1682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 AATCATGGCGTTTGACGACCTTGGAGGTAGACCGGAGTGAAGAACTATGACAAATTCAT 117

QY 1683 GGAAGCAATGGGT 1695
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GGAAGCAATGGGT 130

RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 7218 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
;
; CLONE: pTZgpt-F1s
US-08-232-463-14

```

Query Match 1.8%; Score 43.8; DB 1; Length 7218;
Best Local Similarity 8.4%; Pred. No. 0.37;
Matches 36; Conservative 204; Mismatches 191; Indels 0

	Matches	36;	Conservative	204;	Mismatches	191;	Indels	0;	Gaps	0;
Qy	628	TATCTGTGTCATAAGCAACATTCATTTCAGCATT	CAGACTTAACTAATTTGAAGCTAATA	687	: :	:	: :	:	:	:
Dz	1474	TATCTATGCAAGTAGTTTAAGAAGATAGTGCTACR	RERRERRRRRRRRRRR	1415	: :	:	: :	:	:	:
Qy	688	GACAGCACGATTGGTGC CGTCCATAGGAACA GAC	A CTAACTAATCTCGAGTTAGTACA A	747	: : : : : : :	:	: :	:	:	:
Dz	1414	RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	1355	: : : : : : :	:	: :	:	:	:
Qy	748	GCAGATTAGCAC CAGCAAATTTGCTCAGTTT CA	GTAGTAC TATCTTGTGGGGA GAAG	807	: : : : : : :	:	: :	:	:	:
Dz	1354	RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	1295	: : : : : : :	:	: :	:	:	:

[illegible]

RESIST 10

US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063

```

1 GENERAL INFORMATION:
2 APPLICANT: Dumas Milne Edwards, J.B.
3 APPLICANT: Jobert, S.
4 APPLICANT: Giordano, J.Y.
5 TITLE OF INVENTION: ESTs and Encoded Human Proteins.
6 FILE REFERENCE: GENSET, 054PR2
7 CURRENT APPLICATION NUMBER: US/09/621,976
8 CURRENT FILING DATE: 2000-07-21
9

```

```

; CURRENT FILING DATE: 2000-07-21
;
; NUMBER OF SEQ ID NOS: 19335
;
; SOFTWARE: Patent.pm
;
; SEQ ID NO 2813
;
; LENGTH: 832
;

```

```

;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

```

Query Match 1.8%; Score 43; DB 4; Length 832;
Best Local Similarity 15.0%; Pred. No. 0.21;
Matches 51; Conservative 147; Mismatches 142; Indels

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

RESULT 11

```

US-09-949-016-18019/c
; Sequence 18019, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 18019
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18019

```

Query Match	1.8%	Score 42.6;	DB 4;	Length 601;
Best Local Similarity	54.0%	Pred. No. 0.22;		
Matches 87;	Conservative	0;	Mismatches 74;	Indels 0; Gaps 0;
Qy	2218	AGTCCTCAACTACAGAAAAAGTCACAATAAAAAATGCAAAACATGATGTTCTCTATTGTT	2277	
Db	258	AGTTTCTTCCCACTCAAAATAAACAACAACAACATTAACCTTGGAATAATAACACTTC	199	
Qy	2278	TTTCTCTGCTTGATGTTAATGTAATTAATATTTTTTTTTTTTGGCGGTGAATGTCATGA	2337	
Db	198	CTATGGAGTTGACTTTATTTTCTCCATTGCTTTACCTTTTACAGGTGTTAATAATAGTA	139	
Qy	2338	AAAGAAAGTTAGGAGCCACGATAAATCTGAAGCTCACTATT	2378	
Db	138	AAAGGAAGCTTGAGGTCATGACAAATTTGAAGCTGACAATT	98	

RESULT 12

US-09-949-016-18020/c
; Sequence 18020, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-376

Query Match	1.8%	Score 42.6;	DB 3;	Length 611;
Best Local Similarity	55.0%	Pred. No. 0.23;	59;	Indels 0;
Matches	72;	Conservative 0;	Mismatches 0;	Gaps 0;

Qy	2220	CTCCCTCAACTACAGAAAAAGTCACAATAAAAAATGCAAAACATGATGTTCTATTTTTGT	2279
Db	351	CNCNTTATATTCCNGTTTACCGAAACNAAAAATGGCGNACNTTTTTTTTTTTTTTT	292
Qy	2280	TCCTGCTTGATGTTAATTGATTATTATTTTTTTTTTAGCGGTGAATGTCATGAAA	2339
Db	291	TTTTTTTTTGNNTTTTTTTTTTTTTTTTTTTTTTTTATAAGATTATTAGTATAAAA	232
Qy	2340	AGAAAGTTAGG	2350
Db	231	GGGGAGATAGG	221

Search completed: July 2, 2005, 20:44:57
Job time : 436.909 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 17:53:02 ; Search time 1600.19 Seconds
(without alignments)
9332.708 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agtctctgcgcagaaagg.....atctgaagctcactattcag 2381

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2381	100.0	2381	16	US-10-099-663-1
2	336	14.1	336	16	US-10-099-663-2
3	50.6	2.1	16914	19	US-10-741-601-5698
4	50.6	2.1	16914	21	US-10-741-600-17777
5	49.8	2.1	630	13	US-10-027-632-5694
6	49.8	2.1	630	17	US-10-027-632-5694
7	49.8	2.1	5204	17	US-10-116-275-303

Query Match 100.0%; Score 2381; DB 16; Length 2381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

RESULT 1

US-10-099-663-1
; Sequence 1, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian GUT-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(1626)
; NAME/KEY: exon
; LOCATION: (1627)..(1693)
; NAME/KEY: Intron
; LOCATION: (1694)..(2322)
; NAME/KEY: exon
; LOCATION: (2323)..(2381)
US-10-099-663-1

C	8	47.2	2.0	337	9	US-09-960-352-6036	Sequence 6036, Ap
	9	46.8	2.0	662	20	US-10-335-053-36	Sequence 36, Appl
	10	46.8	2.0	670	21	US-10-764-420-2245	Sequence 2245, Ap
	11	46.2	1.9	6222	15	US-10-311-455-665	Sequence 665, App
	12	44.2	1.9	515	10	US-09-871-161-474	Sequence 474, App
	13	44.2	1.9	2252	19	US-10-741-601-261	Sequence 261, App
	14	44.2	1.9	2252	21	US-10-741-600-745	Sequence 745, Appl
	15	44.2	1.9	3673778	16	US-10-312-841-2	Sequence 2, Appl
	16	43.6	1.8	400	20	US-10-425-115-111538	Sequence 111538
C	17	43.6	1.8	413	19	US-10-674-124A-2159	Sequence 2159, Ap
	18	43.6	1.8	14023	15	US-10-311-455-2077	Sequence 2077, Ap
	19	43.2	1.8	7143	15	US-10-311-455-956	Sequence 956, App
	20	42.8	1.8	51289	19	US-10-322-281-648	Sequence 648, App
C	21	42.6	1.8	513	16	US-10-029-386-13006	Sequence 13006, A
	22	42.6	1.8	611	10	US-09-871-161-376	Sequence 376, App
C	23	42.6	1.8	1241	20	US-10-425-115-124424	Sequence 124424,
	24	42.6	1.8	2844	14	US-10-176-841-19	Sequence 19, Appl
C	25	42.6	1.8	5501	15	US-10-311-455-28	Sequence 28, Appl
	26	42.4	1.8	452	9	US-09-922-217-138	Sequence 138, App
C	27	42.4	1.8	452	9	US-09-833-263-138	Sequence 138, App
C	28	42.4	1.8	452	13	US-10-025-380-138	Sequence 138, App
C	29	42.4	1.8	682	14	US-10-198-846-1721	Sequence 1721, Ap
C	30	42.4	1.8	5908	14	US-10-239-678-94	Sequence 94, Appl
C	31	42.4	1.8	5908	15	US-10-240-453-106	Sequence 106, App
C	32	42.4	1.8	5908	17	US-10-221-613-176	Sequence 176, App
	33	42.2	1.8	453	9	US-09-967-768A-1	Sequence 1, Appl
	34	42.2	1.8	453	10	US-09-873-367C-525	Sequence 525, App
	35	42.2	1.8	453	21	US-10-843-641A-525	Sequence 525, App
	36	42.2	1.8	453	21	US-10-843-641A-6146	Sequence 6146, Ap
C	37	42.2	1.8	783	18	US-10-424-599-83233	Sequence 83233, A
	38	42.2	1.8	6621	15	US-10-311-455-885	Sequence 885, App
	39	42	1.8	9095	19	US-10-433-793-91	Sequence 91, Appl
C	40	41.8	1.8	10048	17	US-10-221-613-211	Sequence 211, App
	41	41.8	1.8	137870	17	US-10-351-951-1	Sequence 1, Appl
C	42	41.6	1.7	624	13	US-10-027-632-194130	Sequence 194130,
C	43	41.6	1.7	624	13	US-10-027-632-194131	Sequence 194131,
C	44	41.6	1.7	624	17	US-10-027-632-194130	Sequence 194130,
C	45	41.6	1.7	624	17	US-10-027-632-194131	Sequence 194131,

QY 1 AGCTTCTCGCGAGAAAAGGCTGTGGGGTCTTTGTTCCCTCTCA CACAGCTTAAGCAAAATC 60
Db 1 AGCTTCTCGCGAGAAAAGGCTGTGGGGTCTTTGTTCCCTCTCA CACAGCTTAAGCAAAATC 60
QY 61 CCCAAGTTCAAAACGCTCGGCTGTAAAGGAGATGGCTCACTTCAAATGAAGTGAATTATG 120
Db 61 CCCAAGTTCAAAACGCTCGGCTGTAAAGGAGATGGCTCACTTCAAATGAAGTGAATTATG 120
QY 121 AAATAATCATAAACAGAGCTCTGTGGCAGATCAGAGATAAACCTCTGCTGGGACAAAAT 180
Db 121 AAATAATCATAAACAGAGCTCTGTGGCAGATCAGAGATAAACCTCTGCTGGGACAAAAT 180
QY 181 CTTAAAGTGTATAGGTAGAAACAGAGAGTGTGGCAACTAAATGGAGCTAGATTTGACACA 240
Db 181 CTTAAAGTGTATAGGTAGAAACAGAGAGTGTGGCAACTAAATGGAGCTAGATTTGACACA 240
QY 241 TTTGATCTTCTAGGAGCAAAAGGGTCTGGAAAACAAAATTAATTTCTGGTGCAAGTCAGT 300
Db 241 TTTGATCTTCTAGGAGCAAAAGGGTCTGGAAAACAAAATTAATTTCTGGTGCAAGTCAGT 300
QY 301 AGCAGCCTGTTTGGGTGCAACTACAGCAACTTTGTTTGCAACAATAACAATCTAAGTTG 360
Db 301 AGCAGCCTGTTTGGGTGCAACTACAGCAACTTTGTTTGCAACAATAACAATCTAAGTTG 360
QY 361 TTTTCTTTCTCTTCTTAACTCTCTGTACAGTCTAAAGGTGAAGAGTACGTATTTGAGT 420
Db 361 TTTTCTTTCTCTTCTTAACTCTCTGTACAGTCTAAAGGTGAAGAGTACGTATTTGAGT 420
QY 421 TACTTCCCTCTGCATCTCTTAGCCAGATTAGCATTTGATTTCAAAATGAACCTGAGTGGA 480
Db 421 TACTTCCCTCTGCATCTCTTAGCCAGATTAGCATTTGATTTCAAAATGAACCTGAGTGGA 480
QY 481 ATGGAAGGCCACTATTTTGGTATCACCAGCAAGTTCTAAATTTATAGTTATACTTCA 540
Db 481 ATGGAAGGCCACTATTTTGGTATCACCAGCAAGTTCTAAATTTATAGTTATACTTCA 540
QY 541 GTAAAACCTTTTGGTGGAGTCTGGAAGAAAAGAAAGATTTATGATAACAGCAGCTAGTA 600
Db 541 GTAAAACCTTTTGGTGGAGTCTGGAAGAAAAGAAAGATTTATGATAACAGCAGCTAGTA 600
QY 601 AAATTCATTTAGTAGCAACCTGTTATCTGTGATAAGCAACATTCATTTACAGCAT 660
Db 601 AAATTCATTTAGTAGCAACCTGTTATCTGTGATAAGCAACATTCATTTACAGCAT 660
QY 661 TCAGGATTTACATTTTGAAGCTAATAGACAGCAGATTTGGTCCGTCCATAGGAACAGA 720
Db 661 TCAGGATTTACATTTTGAAGCTAATAGACAGCAGATTTGGTCCGTCCATAGGAACAGA 720
QY 721 CTAACATATACTGAGTTTGTAGTACAGCAGATTTAGCACCAGCAAAATTTGCTCAGTTTC 780
Db 721 CTAACATATACTGAGTTTGTAGTACAGCAGATTTAGCACCAGCAAAATTTGCTCAGTTTC 780
QY 781 AAGTAGCAGTATCTGTGGGGAAGAGGAGCTCAGCCAGTGTGCTCATTTTCTGCATT 840
Db 781 AAGTAGCAGTATCTGTGGGGAAGAGGAGCTCAGCCAGTGTGCTCATTTTCTGCATT 840
QY 841 ATCCCTCAACATTTAAACCTGGGATCTATGGAATCAAAACAGTGTGGGTAAATTCAC 900
Db 841 ATCCCTCAACATTTAAACCTGGGATCTATGGAATCAAAACAGTGTGGGTAAATTCAC 900
QY 901 TAGCAGCAGATCAACTCTAGGAATGGACAGAAACAGAGCATTTCACTGAATGGGCTAT 960
Db 901 TAGCAGCAGATCAACTCTAGGAATGGACAGAAACAGAGCATTTCACTGAATGGGCTAT 960
QY 961 AATATAGAGAAATACGTAGAGGTGCTCTGAAATTTAGACTACCTATTTAAAGAGTGAGGACA 1020
Db 961 AATATAGAGAAATACGTAGAGGTGCTCTGAAATTTAGACTACCTATTTAAAGAGTGAGGACA 1020
QY 1021 CGAATGGAGAAATATCATCGCAATTTCTGTAGCTCAGCAGTACCTGGAAGGTTTCTGAAA 1080
Db 1021 CGAATGGAGAAATATCATCGCAATTTCTGTAGCTCAGCAGTACCTGGAAGGTTTCTGAAA 1080

QY 1081 CTGAACCGAGTTTCCCAAACTACCTGTGGATGTTTCACTCATCTCATGCTTAT 1140
Db 1081 CTGAACCGAGTTTCCCAAACTACCTGTGGATGTTTCACTCATCTCATGCTTAT 1140
QY 1141 TATGTGGAGTACAAATAGATTTCTCA CCAAAATTAGAATTGGAACAGAGATTTGTGTTTT 1200
Db 1141 TATGTGGAGTACAAATAGATTTCTCA CCAAAATTAGAATTGGAACAGAGATTTGTGTTTT 1200
QY 1201 ATCTGTGGGTAAAATACGTTTTTCTCAGTTGTATAAAGACCTCCCA CCAAGTATAAAGTC 1260
Db 1201 ATCTGTGGGTAAAATACGTTTTTCTCAGTTGTATAAAGACCTCCCA CCAAGTATAAAGTC 1260
QY 1261 CTAATGCAAAAGAAAATGTCATATCTCTCTAGTCTCATTTATTTTTCATTAGATA 1320
Db 1261 CTAATGCAAAAGAAAATGTCATATCTCTCTAGTCTCATTTATTTTTCATTAGATA 1320
QY 1321 GCGGGTTTTTACTCAACTCAAAATAAGATGAACAGAAATGAATGGTTAGTGACTGTTTA 1380
Db 1321 GCGGGTTTTTACTCAACTCAAAATAAGATGAACAGAAATGAATGGTTAGTGACTGTTTA 1380
QY 1381 TAAAGAGAGTAAATAAGATATACTATCATCTTGAAGCAATAAGGAGGAGAGATTGAG 1440
Db 1381 TAAAGAGAGTAAATAAGATATACTATCATCTTGAAGCAATAAGGAGGAGAGATTGAG 1440
QY 1441 CAAAAGTGTGCTTACAGTGGAAAACAGTTTAACTAAAGTGA CCCCCTCCTTGACAA 1500
Db 1441 CAAAAGTGTGCTTACAGTGGAAAACAGTTTAACTAAAGTGA CCCCCTCCTTGACAA 1500
QY 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCATCTAAATTTGCTTTTCTTGATA 1560
Db 1501 GATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCATCTAAATTTGCTTTTCTTGATA 1560
QY 1561 AGCTGTTTCATAAATTTCTTTTGGAAAAGCTCTGCTACTTACCAAGTCTGCTCTACAGAC 1620
Db 1561 AGCTGTTTCATAAATTTCTTTTGGAAAAGCTCTGCTACTTACCAAGTCTGCTCTACAGAC 1620
QY 1621 AGAAGATGGCAATTTAA CCGTACTTTGGAATAAGAGAAAATGAGAACTATCAAAAATTC 1680
Db 1621 AGAAGATGGCAATTTAA CCGTACTTTGGAATAAGAGAAAATGAGAACTATCAAAAATTC 1680
QY 1681 ATGGAAGCAATGGGTAAAGCTTTACTTTTGAATGCTTTCTAAAAGCAGGATACCACTAC 1740
Db 1681 ATGGAAGCAATGGGTAAAGCTTTACTTTTGAATGCTTTCTAAAAGCAGGATACCACTAC 1740
QY 1741 GCGGGAATACAACTTTAAGCTGTTTCA TGAACCTACCATCTGGCTACCTGCTCTTTGTTGT 1800
Db 1741 GCGGGAATACAACTTTAAGCTGTTTCA TGAACCTACCATCTGGCTACCTGCTCTTTGTTGT 1800
QY 1801 CTGCTATTTTGGCTTGGCACTTGCCTGCACTTATTTTGAAGAAAGACTCTATAGAGGGA 1860
Db 1801 CTGCTATTTTGGCTTGGCACTTGCCTGCACTTATTTTGAAGAAAGACTCTATAGAGGGA 1860
QY 1861 ATACAAGGAAGAAAACATTTCTGATTTTATTTTGCATTTGCGATAATCTTATGCAATTTAGCT 1920
Db 1861 ATACAAGGAAGAAAACATTTCTGATTTTATTTTGCATTTGCGATAATCTTATGCAATTTAGCT 1920
QY 1921 AATTCAGTACAGGCAATTCAGAGAAATTTAAATAGAAATTTATATGTAAGGAATATTAT 1980
Db 1921 AATTCAGTACAGGCAATTCAGAGAAATTTAAATAGAAATTTATATGTAAGGAATATTAT 1980
QY 1981 TTGATAAGACTGTTTGAATAATTTACAGAGGGAATTTGCTGGTCTCCAGTTTTCAGACA 2040
Db 1981 TTGATAAGACTGTTTGAATAATTTACAGAGGGAATTTGCTGGTCTCCAGTTTTCAGACA 2040
QY 2041 ACACACATGATTTGAGTCAATTTTAAACATGCTAGTCTTACTTTAAAGCTTTGTACAACTGC 2100
Db 2041 ACACACATGATTTGAGTCAATTTTAAACATGCTAGTCTTACTTTAAAGCTTTGTACAACTGC 2100
QY 2101 CTGTAATATGGATGTAACATTAATCTATCTAGTTGGATAGTAGTTTGTATTAAGGCTCA 2160
Db 2101 CTGTAATATGGATGTAACATTAATCTATCTAGTTGGATAGTAGTTTGTATTAAGGCTCA 2160
QY 2161 ACACTGCCTCAGTGAAGGAGGAGAGTAAAGACTCTGAGTCAGAAATTTCTGGGCTAAGC 2220


```
RESULT 5
US-10-027-632-5694/c
; Sequence 5694, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5694
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5694

Query Match          2.1%; Score 49.8; DB 13; Length 630;
Best Local Similarity 59.2%; Pred. No. 0.09;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCCACATCATGTAAATGCTTCTCGATGAAGCCTGTTCAATAAT 1575
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAGAGTCTGCTACAGACAGAAAGA 1627
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 TCTCGCCCAAGGACAGACCTGAATCTTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 413
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1628 TGGCATTTTAACGGTACTTTGGAAAATAGAGAAAATGAGAAATATGAAAAATTCATGGAAG 1687
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 TGGCGTTTGACAGCACTTGGAGGTAGACCGGAGTGAAAACTATGCAAGTTTCATGGAAG 353
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1688 CAATGGGTAAGCCTTACTTTTGTGAATGCCT 1718
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
US-10-027-632-5694

RESULT 6
US-10-027-632-5694/c
; Sequence 5694, Application US/10027632
; Publication No. US2003020475A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5694
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5694

Query Match          2.1%; Score 49.8; DB 13; Length 630;
Best Local Similarity 59.2%; Pred. No. 0.09;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCCACATCATGTAAATGCTTCTCGATGAAGCCTGTTCAATAAT 1575
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAGAGTCTGCTACAGACAGAAAGA 1627
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 TCTCGCCCAAGGACAGACCTGAATCTTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 413
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1628 TGGCATTTTAACGGTACTTTGGAAAATAGAGAAAATGAGAAATATGAAAAATTCATGGAAG 1687
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 TGGCGTTTGACAGCACTTGGAGGTAGACCGGAGTGAAAACTATGCAAGTTTCATGGAAG 353
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1688 CAATGGGTAAGCCTTACTTTTGTGAATGCCT 1718
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
US-10-027-632-5694

RESULT 7
US-10-116-275-303
; Sequence 303, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; PRIORITY FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 303
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-303

Query Match          2.1%; Score 49.8; DB 17; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.29;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCCACATCATGTAAATGCTTCTCGATGAAGCCTGTTCAATAAT 1575
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 971 TGAACCTTTAAGCTTCCACATCACAGTATGAAGTTGTTTCAAGATAAGAAATATAATAAT 1030
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAGAGTCTGCTACAGACAGAAAGA 1627
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1031 TCTCGCCCAAGGACAGACCTGAATCTTAGCTGCTAGAGGCTGACT-CAACTGAAATCA 1089
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1628 TGGCATTTTAACGGTACTTTGGAAAATAGAGAAAATGAGAAATATGAAAAATTCATGGAAG 1687
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1090 TGGCGTTTGACGACCTTGGAAAGTAGACCGGAGTGAAAACTATGACAAAGTTCATGGAAA 1149
QY 1688 CAATGGGTAAAGCCCTTACTTTTTCGAATGCCT 1718
Db 1150 AATGGGTAAGACTTTTATTTCTTTTGCGCT 1180

RESULT 8
US-09-960-352-6036/c
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagapalan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-K1-G5
US-09-960-352-6036

Query Match 2.0%; Score 47.2; DB 9; Length 337;
Best Local Similarity 53.9%; Pred. No. 0.29;
Matches 97; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 2185 AAGAGTAAGACTCTGACTGAGATTCGGGCTAAGCTCCCTCACTACAGAAAAAGTCAC 2244
Db 289 AATAAATATAATATATTTAAACAAATTAATAAGTTACATATATATAAAAAAAATGTGAT 230
QY 2245 AATAAAAAATGCAACATGATGTTCTATTTTGTGTTTCTCTGCTTGATGTTAATTGATTAT 2304
Db 229 ATTATATATTAAATATT 170
QY 2305 TATATATTTTTTTTTTTAGGCGTGAATGTGATGAAAAAGAAAGTTAGGAGCCACGATAATC 2364
Db 169 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATTAAGTGTATTACTTTATTTCAATATC 110

RESULT 9
US-10-335-053-36
; Sequence 36, Application US/10335053
; Publication No. US20040241653A1
; GENERAL INFORMATION:
; APPLICANT: Quark Biotech, Inc.
; TITLE OF INVENTION: Methods for identifying marker genes for cancer
; FILE REFERENCE: 68733-A; 070/US1
; CURRENT APPLICATION NUMBER: US/10/335.053
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: 60/345,317
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-335-053-36

Query Match 2.0%; Score 46.8; DB 20; Length 662;
Best Local Similarity 73.2%; Pred. No. 0.53;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1615 ACAGACAGAAGATGCGTATTAAACGGTACTTGGAAATAGAGAAAAATGAGAACTATGAA 1674

Db 28 ACAGCTGAGATCATGCGCATTCGACGCGACCTGGAAAAGTAGACCGGACGAGAACTATGAA 87
QY 1675 AAATTCATGGAAGCAATGGGTA 1696
Db 88 AAGTTCATGGAGAAATGGGCA 109

RESULT 10
US-10-764-420-2245
; Sequence 2245, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Dai, Hongyue
; APPLICANT: Tan, Yejun
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2245
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-2245

Query Match 2.0%; Score 46.8; DB 21; Length 670;
Best Local Similarity 73.2%; Pred. No. 0.53;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1615 ACAGACAGAAGATGCGTATTAAACGGTACTTGGAAAATAGAGAAAAATGAGAACTATGAA 1674
Db 24 ACAGCTGAGATCATGCGCATTCGACGCGCACCTGGAAAAGTAGACCGGAAACGAGAACTATGAA 83
QY 1675 AAATTCATGGAAGCAATGGGTA 1696
Db 84 AAGTTCATGGAGAAATGGGCA 105

RESULT 11
US-10-311-455-665
; Sequence 665, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 665
; LENGTH: 6222
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-665

Query Match 1.9%; Score 46.2; DB 15; Length 6222;
Best Local Similarity 66.7%; Pred. No. 2.6; Mismatches 0; Gaps 0;
Matches 66; Conservative 0; Indels 33; Indels 0; Gaps 0;
QY 1623 AAAGATGGCATTAAACGGTACTTTGGAATAATAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 4958 AATATGGCGTTTCATAGTATTTGGAAGGTAGATCGAGTGAATAATTAATGATAAGTTTAT 5017
QY 1683 GGAAGCAATGGTAAAGCCTTACTTTTTTGAATGCCCTTCT 1721
DB 5018 GGAAAAAATGGGTAAAGATTTTATTTTTTTTGIGGTTTAT 5056

RESULT 12
US-09-871-161-474
; Sequence 474, Application US/09871161
; Publication No. US20030097666A1
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/871,161
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/328,111
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/117,393
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 60/098,639
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 474
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(515)
; OTHER INFORMATION: n = A,T,C or G
US-09-871-161-474

Query Match 1.9%; Score 44.2; DB 10; Length 515;
Best Local Similarity 75.3%; Pred. No. 2.1; Mismatches 0; Gaps 0;
Matches 55; Conservative 0; Indels 18; Indels 0; Gaps 0;
QY 1623 AAAGATGGCATTAAACGGTACTTTGGAATAATAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 23 AATCATGGCGTTTGACAGCACTTTGGAAGGTAGACCGGAGTGAATACTATGACAAGTTTCAT 82
QY 1683 GGAAGCAATGGGT 1695
DB 83 GGAAAAAATGGGT 95

RESULT 13
US-10-741-601-261
; Sequence 261, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-261

Query Match 1.9%; Score 44.2; DB 19; Length 2252;
Best Local Similarity 75.3%; Pred. No. 4.7; Mismatches 0; Gaps 0;
Matches 55; Conservative 0; Indels 18; Indels 0; Gaps 0;
QY 1623 AAAGATGGCATTAAACGGTACTTTGGAATAATAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCACTTTGGAAGGTAGACCGGAGTGAATACTATGACAAGTTTCAT 117
QY 1683 GGAAGCAATGGGT 1695
DB 118 GGAAAAAATGGGT 130

RESULT 14
US-10-741-600-745
; Sequence 745, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 745
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-745

Query Match 1.9%; Score 44.2; DB 21; Length 2252;
Best Local Similarity 75.3%; Pred. No. 4.7; Mismatches 0; Gaps 0;
Matches 55; Conservative 0; Indels 18; Indels 0; Gaps 0;
QY 1623 AAAGATGGCATTAAACGGTACTTTGGAATAATAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
DB 58 AATCATGGCGTTTGACAGCACTTTGGAAGGTAGACCGGAGTGAATACTATGACAAGTTTCAT 117
QY 1683 GGAAGCAATGGGT 1695
DB 118 GGAAAAAATGGGT 130

RESULT 15
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MRNOMPHISMS ASSOCIATED WITH STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 1.9%; Score 44.2; DB 16; Length 3673778;
Best Local Similarity 46.1%; Pred. No. 2.6e+02; Mismatches 183; Conservative 0; Indels 213; Indels 1; Gaps 1;
QY 1951 TAAATAGAATTATATCTAAGGAATATTATTTTGTATAAGACTCTTTTGAATACTATGACAAG 2010

```
Db 3262845 TGAAGGGAAGAGGTTAATGGAGGCGAGTTGGATTTTATTTGTAGAAAAATGGGAATA 3262904
Qy 2011 AGGGAATTTGCTGCTCCAGTTTTCAGAACACACATGATTTTGAGTCATTTTAAACATGC 2070
Db 3262905 ATTGAAGGTGTAGTTTTTTTTTTTAAAGTAATATTTATTTGGGATTTTTTAAATTGTAA 3262964
Qy 2071 TAGTGCTTACTTTAAGCTTGTACAACTGGCTGTAATATGGATGTAAACATAACTATCCT 2130
Db 3262965 AAGTAAATATGTTAATTGTTGTATTTGGAAAGTATAGAAAAGTATAAAAGATATAAT 3263024
Qy 2131 AGTTGGATAGTAGTTTGTATACAGGCTGAACACTGCCTCAGTGAAAGGTGGAGAAGT 2190
Db 3263025 TATTTATAATTTTATTAATTAATTAGTAATAATAATTATTTGGTATTTTATTTGTTTGT 3263084
Qy 2191 AAGACTCTGAGTCAGAAATTCGGGCTAAGCTCCCTCAACTACAGAAAAAGTCACAATAAA 2250
Db 3263085 AAGTTGATGTTAGTTTTTTTATTTTAAAGTTTATTTTGAAGTTTATTCG-AAAATTGGA 3263143
Qy 2251 AATGCAACATGATGTTCTATTTTGTGTTTTTCTCTGCTTGATGTTAAATTGATTATTTAT 2310
Db 3263144 ATTATACGATGTTGTTTTTATATTTTTTTTGTGTTTTTATTTTTTTTTTTTATAATATTTAT 3263203
Qy 2311 TTTTTTTTTTAGGCGTGAATGTGATGAAAGAGTT 2347
Db 3263204 TTTTTTTTGAATTTTAAATGTGAAAAAAAATTTT 3263240
```

Search completed: July 3, 2005, 00:59:42
Job time : 1613.19 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 13:41:36 ; Search time 8614.37 Seconds
(without alignments)
10520.918 Million cell updates/sec

Title: US-10-099-663-1
Perfect score: 2381
Sequence: 1 agctctgcgcagaaag.....atctgaagctcactattcag 2381

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	752.6	31.6	1473	8	CC250935 CH261-17C
2	99	4.2	598	7	CF250818 esa017_e0
3	95	4.0	595	7	CF250947 esa019_a0
4	85	3.6	705	5	BU355106 603474288
5	84	3.5	759	5	BU265304 603508640
6	81	3.4	829	5	BU296697 603741965
7	79.4	3.3	885	5	BU123280 603003274
8	79	3.3	814	5	BU234099 603792609
9	78.8	3.3	1080	5	BU123336 603149475
10	75.8	3.2	426	6	CD739695 4028031_1
11	74	3.1	576	6	CD739141 4026481_1
12	70.4	3.0	480	6	CD733346 4045132_1
C 13	62.4	2.6	1101	9	CNS0039G
14	61.4	2.6	885	5	BU123717 603147391
15	60.4	2.5	443	5	EX276255 BX276255
16	57.2	2.4	683	7	CN793516 4128553_B
17	55.8	2.3	657	7	CK945920 4070423_B
18	54.2	2.3	572	7	CK834687 4059069_B
19	54.2	2.3	673	7	CK947902 4072726_B
20	54.2	2.3	686	7	CK957350 4097834_B
21	54.2	2.3	708	7	CK960577 4101917_B
22	54	2.3	481	4	BM430667 1DUO3A11
23	54	2.3	489	6	CB224035 1JEU30F5
24	54	2.3	497	4	BM430956 1DUO07F03

25	54	2.3	548	4	BM432416
26	54	2.3	568	4	BM431104
27	54	2.3	671	7	CK980160
28	54	2.3	679	7	CK947198
29	54	2.3	682	7	CK948900
C 30	54	2.3	684	7	CK947048
31	54	2.3	702	7	CK944155
32	54	2.3	720	7	CK946512
33	53	2.2	372	4	BM430450
34	53	2.2	414	4	BM433066
35	53	2.2	416	4	BM432621
36	53	2.2	546	7	CN098758
C 37	53	2.2	576	7	CN098757
38	52.6	2.2	660	7	CK954372
39	52.6	2.2	1101	9	CNS0039G
40	52.4	2.2	503	4	BM432863
41	52	2.2	339	4	BM430492
42	52	2.2	371	4	BM432453
43	52	2.2	371	4	BM433013
44	52	2.2	415	4	BM430284
45	52	2.2	433	4	BM431869

ALIGNMENTS

RESULT 1
CC250935/c
LOCUS CH261-17C7.Sp6.1 CH261 Gallus gallus genomic clone CH261-17C7, 1473 bp DNA linear GSS 13-MAY-2003
DEFINITION genomic survey sequence.
ACCESSION CC250935
VERSION CC250935.1 GI:30587685
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1473)
AUTHORS Kramitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 239
High quality sequence stop: 912.
Location/Qualifiers
1. .1473
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-17C7"
/sex="female"
/cell_line="UCD001, inbred 256"
/notes="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac"

FEATURES

source

ORIGIN

Query Match 31.6%; Score 752.6; DB 8; Length 1473;
Best Local Similarity 98.3%; Pred. No. 4.3e-168;
Matches 793; Conservative 0; Mismatches 9; Indels 5; Gaps 3;
QY 1580 TTTCGAAGCTCTGCTACTTACCAGAAGTC-TGCCTACAGACAGAAAGATGGCATTTAA- 1637

Db	1042	TTGCAAAAGTTCCTGCTTCTTACCCAGAAATCTTCCCTACAGACAGAAAGATGGCACTTTTAAAC	983
Qy	1638	CGGTACTTTGGAAAAATAGAGAAAAATAGAACTATGAAAAATTCATGGAAGCAATGGGTAA	1697
Db	982	CGGTACTTTGGAAAAATAGAGAAAAATAGAACTATGAAAAATTCATGGAAGCAATGGGTAA	923
Qy	1698	GCCTTACTTTTTTGAATGCCCTTTCAAAGCAGAGATACACCTACGCGCGGAATACAAACTTA	1757
Db	922	GCCTTACTTTTTTGAATGCCCTTTCAAAGCAGAGATACACCTACGCGCGGAATACAAACTTA	863
Qy	1758	AGCTGTTTCATGAACCTACCATCTGCTTAACCTGCTTTGTTGTCGTCTATTTTGGCCCTTG	1817
Db	862	AGCTGTTTCATGAACCTACCACTGCTTAACCTGCTTTGTTGTCGTCTATTTTGGCCCTTG	803
Qy	1818	CACATTGCCCTGCACTTATTTTGGAAAAAGACTCTATAGAGGGGAATACAAGGAAGAAAAAC	1877
Db	802	CACATTGCCCTGCACTTATTTTGGAAAAAGACTCTATAGAGGGGAATACAAGGAAGAAAAAC	743
Qy	1878	ATTCTGATTTTTTATTTGGCATTCGGATATCTTATGCATTTAGCTAATTTCCAGTAGAGGCAT	1937
Db	742	ATTCTGATTTTTTATTTGGCATTCGGATATCTTATGCATTTAGCTAATTTCCAGTAGAGGCAT	683
Qy	1938	TCCAGCAGAAATTTAAATAGAAATATATGTAAGGAAATATTTTGTGATAAGACTGTTTGA	1997
Db	682	TCCAGCAGAAATTTAAATAGAAATATATGTAAGGAAATATTTTGTGATAAGACTGTTTGA	623
Qy	1998	AAAAATTACAGAGGGGAAATTCGTGCTCCTCAGTTTTTGCAGAACACACATGATTTGAGT	2057
Db	622	AAAAATTACAGAGGGGAAATTCGTGCTCCTCAGTTTTTGCAGAACACACATGATTTGAGT	563
Qy	2058	CATTTTAAACATGCTAGTGCTTTACTTTTAAGCTTTGACAAACTGCTGCTTAATGATGCTAA	2117
Db	562	CATTTTAAACATGCTAATGCTTTACTTTTAAGCTTTGACAAACTGCTGCTTAATGATGCTAA	503
Qy	2118	ACATAACTATCCTTAGTTGGATAGTAGTTGTATTTACAGGCTGAACACTGCCTCAGTGAAA	2177
Db	502	ACATAACTATCCTTAGTTGGATAGTAGTTGTATTTACAGGCTGAACACTGCCTCAGTGAAA	443
Qy	2178	GGTGGAGAGAGTAAGACTCTGAGTCAGAAATCTGGGCTAAGCTCCCTCAACTACAGAAA	2237
Db	442	GGTGGAGAGAGTAAGACTCTGAGTCAGAAATCTGGGCTAAGCTCCCTCAACTACAGAAA	383
Qy	2238	AAGTCACAATAAAAAATCAACATGATGTTCTATTTTGTGTTTTTCTCTGCTTGATGTTAAT	2297
Db	382	AAGTCACAATAAAAAATCAACATGATGTTCTATTTTGTGTTTTTCTCTGCTTGATGTTAAT	323
Qy	2298	TGATTATATTA---TTTTTTTTTTTAGGGGTGAATGATGATAAGAAAGAAAGTTAGAGCC	2354
Db	322	TGATTATATTAATTTTTTTTTTTTTTAGGGGTGAATGATGATAAGAAAGAAAGTTAGAGCC	263
Qy	2355	CACGATAATCTGAAGCTCACATTTCAG	2381
Db	262	CACGATAATCTGAAGCTCACATTTCAG	236
RESULT 2			
CF250818			
LOCUS			
DEFINITION	esao17 e05 Elmeiria tenella-infected caecal tonsil Gallus gallus	598 bp	linear EST 07-AUG-2003
ACCESSION	CF250818		
VERSION	CF250818.1	GI:33484073	
KEYWORDS	EST.		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;		
	Phasianinae; Gallus.		
	1. (bases 1 to 598)		
	Witzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whittaker,C.J.,		
	Chausse,A.M. and Zoorob,R.		
TITLE	A collection of chicken ESTs from activated immune cells		

```

Db      63  AATGAGAACTATGAAAAATTCATGGAAGCAATGGG 97
|||||
RESULT 4
BU355106
LOCUS   603474288F1 CSEQCHN70 Gallus gallus cDNA clone CHEST355022 5', mRNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 705)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
PUBMED
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .705
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST355022"
/dev_stage="36"
/lab_host="DH108"
/clone_lib="CSEQCHN70"
/notes="Organ: hearts; Vector: pBluescript II KS(+); Site 1:
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
ORIGIN
Query Match 3.6%; Score 85; DB 5; Length 705;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1610 TGCTACAGACAGAAAGATGCGATTTTAAACGGTACTTGGAAATAGAGAAAAATGAGAAGT 1669
|||||
Db 1 TGCTACAGACAGAAAGATGCGATTTTAAACGGTACTTGGAAATAGAGAAAAATGAGAAGT 60
QY 1670 ATGAAAAATTCATGGAAGCAATGGG 1694
|||||
Db 61 ATGAAAAATTCATGGAAGCAATGGG 85
|||||
RESULT 5
BU265304
LOCUS   603508640F1 CSEQCHN52 Gallus gallus cDNA clone CHEST437010 5', mRNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 759)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
PUBMED
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .759
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST437010"
/dev_stage="22"
/lab_host="DH108"
/clone_lib="CSEQCHN52"
/notes="Organ: limbs; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
ORIGIN
Query Match 3.5%; Score 84; DB 5; Length 759;
Best Local Similarity 100.0%; Pred. No. 7.4e-09;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1611 GCCTACAGACAGAAAGATGCGATTTTAAACGGTACTTGGAAATAGAGAAAAATGAGAAGT 1670
|||||
Db 1 GCCTACAGACAGAAAGATGCGATTTTAAACGGTACTTGGAAATAGAGAAAAATGAGAAGT 60
QY 1671 TGAATAATTCATGGAAGCAATGGG 1694
|||||
Db 61 TGAATAATTCATGGAAGCAATGGG 84
|||||
RESULT 6
BU296697
LOCUS   603741965F1 CSEQCHN56 Gallus gallus cDNA clone CHEST640b17 5', mRNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
EST.

```

```

SOURCE          Gallus gallus (chicken)
ORGANISM         Gallus gallus
REFERENCE        1 (bases 1 to 829)
AUTHORS          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                  Phasianinae; Gallus.
TITLE            1 (bases 1 to 829)
JOURNAL          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
MEDLINE          Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED          22335534
COMMENT          12445392
                  Contact: Simon Hubbard
                  Department of Biomolecular Sciences
                  University of Manchester Institute of Science and Technology
                  (UMIST)
                  PO Box 88, Manchester, M60 1QD, UK
                  Tel: 01612008930
                  Fax: 01612360409
                  Email: Simon.Hubbard@umist.ac.uk.

FEATURES
  source
    1..829
      /organism="Gallus gallus"
      /mol_type="mRNA"
      /strain="Compton Line 151"
      /db_xref="taxon:9031"
      /clone="CHEST14c20"
      /sex="Female"
      /dev_stage="adult"
      /lab_host="DH10B"
      /clone_lib="CSEQCHN56"
      /note="Organ: small intestine; Vector: pBluescript II
      KS(+); Site 1: EcoRI; Site 2: NotI; This normalized
      library was constructed from 1 million independent clones.
      cDNA synthesis was initiated using an oligo(dT) primer,
      using methylated C in the first strand synthesis reaction.
      Following this first strand reaction, double-stranded cDNA
      was bluntended, ligated to NotI adapters, digested with
      EcoRI, size-selected, and cloned into the NotI and EcoRI
      compatible sites of a custom modified MCS of the
      pBluescript (KS+) vector. The library was normalized in 2
      rounds using conditions adapted from Soares et al., PNAS
      (1994) 91: 9228-9232 and Bonaldo et al., Genome Research
      (1996) 791, except that a significantly longer
      reannealing hybridization was used."

ORIGIN
Query Match      3.4%; Score 81; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1614 TACAGACAGAAAGATGCGCATTTAACGGTACTTGGAAAAATAGAGAAAAATGAGAACTATGA 1673
Db 11 TACAGACAGAAAGATGCGCATTTAACGGTACTTGGAAAAATAGAGAAAAATGAGAACTATGA 70

QY 1674 AAAATTCATGGAAGCAATGGG 1694
Db 71 AAAATTCATGGAAGCAATGGG 91

RESULT 7
BU123280
LOCUS 603003274F1 CSEQCHL18 Gallus gallus cDNA clone CHEST14c20 5', mRNA
DEFINITION sequence.
ACCESSION BU123280
VERSION BU123280.1 GI:25333903
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE        1 (bases 1 to 814)
AUTHORS          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
TITLE            1 (bases 1 to 814)
JOURNAL          Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE          22335534
PUBMED          12445392
                  Contact: Simon Hubbard

SOURCE          Gallus gallus (chicken)
ORGANISM         Gallus gallus
REFERENCE        1 (bases 1 to 895)
AUTHORS          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
TITLE            1 (bases 1 to 895)
JOURNAL          Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE          22335534
PUBMED          12445392
COMMENT          Contact: Simon Hubbard
                  Department of Biomolecular Sciences
                  University of Manchester Institute of Science and Technology
                  (UMIST)
                  PO Box 88, Manchester, M60 1QD, UK
                  Tel: 01612008930
                  Fax: 01612360409
                  Email: Simon.Hubbard@umist.ac.uk.

FEATURES
  source
    1..895
      /organism="Gallus gallus"
      /mol_type="mRNA"
      /strain="Compton Line 151"
      /db_xref="taxon:9031"
      /clone="CHEST14c20"
      /sex="Female"
      /dev_stage="adult"
      /lab_host="DH10B"
      /clone_lib="CSEQCHL18"
      /note="Organ: small intestine; Vector: pBluescript II
      KS(+); Site 1: EcoRI; Site 2: NotI; Modification of
      pBluescript II KS(+) [Stratagene] vector to accommodate
      cDNA produced with the T-trimmed protocol (Construcation of
      uni-directionally cloned cDNA libraries from messenger RNA
      for improved 3' end DNA sequencing by Glenn Fu, et al.
      U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with
      NotI and EcoRI. Ligate in double stranded adaptor
      containing BsgI and BamHI sites
      [5'gccgcgctgcagcccgatccgaaacaaag]
      [5'aatttttttcggatccgggctgcagcgc]"

ORIGIN
Query Match      3.3%; Score 79.4; DB 5; Length 895;
Best Local Similarity 98.8%; Pred. No. 9.6e-08;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1614 TACAGACAGAAAGATGCGCATTTAACGGTACTTGGAAAAATAGAGAAAAATGAGAACTATGA 1673
Db 1 TACAGACAGAAAGATGCGCATTTAACGGTACTTGGAAAAATAGAGAAAAATGAGAACTATGA 60

QY 1674 AAAATTCATGGAAGCAATGGG 1694
Db 61 AAAATTCATGGAAGCAATGGG 81

RESULT 8
BU234099
LOCUS 603792609F1 CSEQCHN24 Gallus gallus cDNA clone CHEST758b4 5', mRNA
DEFINITION sequence.
ACCESSION BU234099
VERSION BU234099.1 GI:25478348
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE        1 (bases 1 to 814)
AUTHORS          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
TITLE            1 (bases 1 to 814)
JOURNAL          Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE          22335534
PUBMED          12445392
                  Contact: Simon Hubbard

```

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

FEATURES

Source

```

/organism="Callus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChST758b4"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSEQCIN24"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

```

Email: Simon.Hubbard@umist.ac.uk

ORIGIN

Query Match	3.3%;	Score 78.8;	DB 5;	Length 1080;
Best Local Similarity	97.6%;			
Pred. No.	1.4e-07;			
Matches	80;			
Conservative				
Mismatches	2;			
Indels	0;			

FEATURES

Source

```

/mol_type="mRNA"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_20B02"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="IGAL - Chicken Intestinal Lymphocyte"
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:
Sali; Site 2: NotI; Normalized library from chicken gut
infected with coccidia duodenum and middle gut."

ORIGIN
Query Match          3.2%; Score 75.8; DB 6; Length 426;
Best Local Similarity 97.5%; Pred. No. 5.8e-07;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1616 CAGACAGAAAGATGGCATTTAACGGTACTTGGAAATAGAGAAAAATGAGAACTATGAAA 1675
      |||||
Db 2 CAGACAGAAAGATGGCGTTTAAACGGTACTTGGAAATAGAGAAAAATGAGAACTATGAAA 61
      |||||

QY 1676 AATTCATGAAGCAATGGG 1694
      |||||
Db 62 AATTCATGAAGCAATGGG 80
      |||||

RESULT 11
CD739141
LOCUS
DEFINITION
4026481 IGAL - Chicken Intestinal Lymphocyte Gallus gallus cDNA
CD739141
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 576)
AUTHORS
Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
Tassel,C. and Han,J.Y.
TITLE
Chicken intestinal lymphocyte EST database as a resource for the
analysis of mucosal immune function
JOURNAL
Unpublished (2003)
COMMENT
Contact: Hyun S. Lillehoj
Animal Parasite Diseases Laboratory
Animal and Natural Resources Institute, USDA
Bldg.1043, BARC-East, Beltsville, MD 20705, USA
Tel: 3015048771
Fax: 3015045103
Email: hlilleho@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim_alt '- -trim_fasta. Vector identified
by cross match using options -minmatch 12 -minscore 12
Plate: 93 row: A column: 12
Seq primer: ATTAGGTGACACTATAG
High quality sequence stop: 576.
Location/Qualifiers
1..576
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_93A12"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="IGAL - Chicken Intestinal Lymphocyte"

FEATURES
source
1..576
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_93A12"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="IGAL - Chicken Intestinal Lymphocyte"

/mol_type="mRNA"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_20B02"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="IGAL - Chicken Intestinal Lymphocyte"
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site 1:
Sali; Site 2: NotI; Normalized library from chicken gut
infected with coccidia duodenum and middle gut."

ORIGIN
Query Match          3.0%; Score 70.4; DB 6; Length 480;
Best Local Similarity 98.6%; Pred. No. 1.2e-05;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1623 AAAGATGGCATTTAACGGTACTTGGAAATAGAGAAAAATGAGAACTATGAAAAATTCAT 1682
      |||||

```


Query Match 2.6%; Score 61.4; DB 5; Length 885;
Best Local Similarity 96.1%; Pred. No. 0.0019;
Matches 74; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1618 GACAGAAAGATGCGATTACGGTACTTGGAAAAATAGAGAAAAATGAGAACTATGAAAAA 1677
|||||
Db 8 GAGAGAAAGATGCGATTACGGTACTTGGAAAAAT--AGAAAAATGAGAACTATGAAAAA 65
|||||

QY 1678 TTCATGGAAGCAATGGG 1694
|||||
Db 66 TTCATGGAAGCAATGGG 82
|||||

RESULT 15
BX276255 443 bp mRNA linear EST 15-JUL-2004
LOCUS
DEFINITION
BX276255 AGENAE Gallus gallus multi-tissues normalized library
(gcag) Gallus gallus cdna clone gcag0008c.c.24 5prim, mRNA
sequence.
ACCESSION
BX276255
VERSION
BX276255.1 GI:28598746
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 443)
Herault,F., Le Meuth-Metzinger,V., Desert,C., Retout,E., Piumi,F.,
Klopp,C. and Douaire M.
AUTHORS
Construction and Primary characterization of chicken normalized
multi-tissue cDNA libraries
JOURNAL
Unpublished (2003)
COMMENT
Contact: Douaire M
INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63
Fax: +33 (0) 2.23.48.54.70
Email: Madeleine.Douaire@roazhon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0008 row: c column: 24
Seq primer: M13R.

FEATURES
source
1..443
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcag0008c.c.24"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized
library (gcag)"
/notes="Vector: pT73D-pac; tissues: brain, embryos,
kidney, multi-tissues, muscle, pancreas, skin, testis,
liver, adipose tissue, granulosa, utero-vaginal gland,
oviduct, small follicle, ovary, hypothalamus, pituitary
gland, ileon, jejunum, caecum, duodenum, spleen,
fabricius gland, bone marrow, thymus, hematopoietic
progenitor cells. Clone distribution : AGENAE Resource
centre, Francois PIUMI, Francois.Piumi.inra.fr, INRA, CEA
Radiobiologie et Etude du genome (LREG), Domaine de
Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN
Query Match 2.5%; Score 60.4; DB 5; Length 443;
Best Local Similarity 98.4%; Pred. No. 0.0027;
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2320 TAGGCGTGAATGTGATGAAAAAGTTAGGAGCCACGATAATCTGAAGCTCACTATTTC 2379
|||||

Db 31 TGGGCGTGAATGTGATGAAAAAGTTAGGAGCCACGATAATCTGAAGCTCACTATTTC 90
QY 2380 AG 2381
||
Db 91 AG 92

Search completed: July 2, 2005, 20:36:38
Job time : 8620.37 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length				
C 1	42.8	12.7	88688	9	AC087429	AC087429 Homo sapi	
C 2	42.8	12.7	187003	2	AC023353	AC023353 Homo sapi	
C 3	41.8	12.4	252420	3	AE014841	AE014841 Plasmodiu	
C 4	41.4	12.3	1663	3	AK116155	AK116155 Clona int	
C 5	41.4	12.3	98855	8	OSJN00175	OSJN00175	
C 6	41.4	12.3	121130	8	OSJN00292	AL162973 Oryza sat	
C 7	40.6	12.1	209024	10	AL928605	BX548156 Oryza sat	
C 8	40.4	12.0	60755	2	AL583823	AL928605 Mouse DNA	
C 9	40.4	12.0	131708	9	AL354710	Continuation (4 of	
C 10	40.4	12.0	174294	9	AC026155	AL354710 Human DNA	
C 11	40.2	12.0	110000	10	AE014175	AC026155 Homo sapi	
C 12	40.2	12.0	201783	10	AC114410	Continuation (3 of	
C 13	39.8	11.8	177655	2	CR318618	AC114410 Mus muscu	
C 14	39.8	11.8	217375	5	BX537337	CR318618 Danio rer	
C 15	39.6	11.8	162605	2	AC068698	BX537337 Zebrafish	
C 16	39.6	11.8	164275	9	AC005157	AC068698 Homo sapi	
C 17	39.6	11.8	230585	10	AC110259	AC005157 Homo sapi	
C 18	39.4	11.7	393	8	YMC0R15	AC110259 Mus muscu	
C 19	39.4	11.7	396	8	MIECOR15	K02492 Yeast (S.c	
C 20	39.4	11.7	396	8	MIECOR15	K02126 Yeast mitoc	

Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.
Direct Submission
Submitted (11-APR-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On Apr 11, 2002 this sequence version replaced gi:12039240.

TITLE
JOURNAL

COMMENT

-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hg.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgc@igtp.ac.cn
----- Project Information
Center project name:18 project
Center clone name: 2063K18
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 5% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89950 bases at least Q40
Consensus quality: 90370 bases at least Q30
Consensus quality: 90443 bases at least Q20
Insert size: 88688; sum-of-contigs
Quality coverage: 7.47x in Q20 bases;sum-of-contigs

----- Location/Qualifiers
source
1. .88688
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="2063K18"

FEATURES
source

----- Location/Qualifiers
source
1. .88688
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="2063K18"

ORIGIN

Query Match 12.7%; Score 42.8; DB 9; Length 88688;
Best Local Similarity 53.6%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 ATTATTATTTCATTAGATAGCGGTTTTTACTCAACCTCAATTAAGATGAACAATG 60
DB 1637 ATTACATTAACAATAATTAGTAGGCTTTATATATCTTTTAACACGAACTTAGAAGACT 1578
QY 61 AATGGGTAGTACTGTTTATAAAGAGATTAAGAATACATCATCATCATTTGAGGCAA 120
DB 1577 TAAGAGTAAGTCCATGTATATAAAGATAAGATAAGACACTATATAAGAATAAGCAA 1518
QY 121 TAAGGGAGGAGAGATTACGAACAGTGTGCTTACAAGTGGNAAA 166
DB 1517 TAAGAACACTAAATACTTAGAAATAAATTTAATAAAAATGTACAAA 1472

RESULT 2
AC023353/c
LOCUS Homo sapiens chromosome 20 clone RP11-775C23, WORKING DRAFT
DEFINITION SEQUENCE, 28 unordered pieces.
ACCESSION AC023353
VERSION AC023353.3 GI:9211518
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 187003)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187003)
AUTHORS Waterston,R.H.
TITLE Direct Submission

JOURNAL

COMMENT

Submitted (11-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 15, 2000 this sequence version replaced gi:7235346.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information
Center project name: H.NH0775C23
----- Summary Statistics
Sequencing vector: M13; 57%
Sequencing vector: plasmid; 43%
Chemistry: Dye-primer ET; 57% of reads
Chemistry: Dye-terminator Big Dye; 43% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 172430 bases at least Q40
Consensus quality: 176603 bases at least Q30
Insert size: 189000; agarose-fp
Insert size: 184303; sum-of-contigs
Quality coverage: 3.50 in Q20 bases; agarose-fp
Quality coverage: 3.66 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1408: contig of 1408 bp in length
1409 1508: gap of unknown length
1509 3342: contig of 1834 bp in length
3343 3442: gap of unknown length
3443 5329: contig of 1887 bp in length
5330 5429: gap of unknown length
5430 8611: contig of 3182 bp in length
8612 8711: gap of unknown length
8712 11044: contig of 2333 bp in length
11045 11144: gap of unknown length
11145 15105: contig of 3961 bp in length
15106 15205: gap of unknown length
15206 17732: contig of 2527 bp in length
17733 17832: gap of unknown length
17833 21791: contig of 3959 bp in length
21792 21891: gap of unknown length
21892 23477: contig of 1656 bp in length
23478 23647: gap of unknown length
23648 28121: contig of 4474 bp in length
28122 28221: gap of unknown length
28222 34018: contig of 5797 bp in length
34019 34118: gap of unknown length
34119 38469: contig of 4351 bp in length
38470 38569: gap of unknown length
38570 44271: contig of 5702 bp in length
44272 44371: gap of unknown length
44372 51457: contig of 7086 bp in length
51458 51557: gap of unknown length
51558 58379: contig of 7422 bp in length
58380 59079: gap of unknown length
59080 65759: contig of 6580 bp in length
65760 65759: gap of unknown length
65760 74077: contig of 8318 bp in length
74078 74177: gap of unknown length
74178 82579: contig of 8402 bp in length
82580 82679: gap of unknown length
82680 92223: contig of 9544 bp in length
92224 92323: gap of unknown length
92324 99655: contig of 7332 bp in length
99656 99755: gap of unknown length


```
/note="AT_rich"
/rpt_type=tandem
repeat_region 6419. .6479
/rpt_type=tandem
repeat_region complement(6419. .6479)
/rpt_type=tandem
/rpt_unit="(ta)n"
complement(6497. .6686)
/note="AT_rich"

Query Match 12.4%; Score 41.8; DB 3; Length 252420;
Best Local Similarity 50.2%; Pred. No. 2.1; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 102;

Qy 14 TTGATAGCCGGTTTTTTTACTACAACTCAATAAGATGAACAGAGTAAGTGGGTAGTGA 73
Db 77046 TTAATTATGATGAATATATAGAAATGGAATAAATAAATAAATAAAGGAAGATCAATAA 77105

Qy 74 CTGTTTATAAGAGAGTAATAAGATCACTATCATTTTGGGCAATAGGGGAGAGA 133
Db 77106 ATGTCCTTTTATAAAGAAATCAAAATGAGGAAGATTCTTCTACAAAAAGTTGAATAGTA 77165

Qy 134 GATTTCAGCAACAGCTGTCTTACAAAGTGGAAACAAAGTTAAACTTAAAGTGACCCCTCC 193
Db 77166 CAAGTAAATAAAGCTGTGCAACAAAGGAAACAAAAAAGAGATGTTATCACAAAGG 77225

Qy 194 TTGACAAAGATCAATGCCACAGTTGA 218
Db 77226 TACATAAACAACAACAAATTTGA 77250

RESULT 4
AK116155/c
LOCUS AK116155 1663 bp mRNA linear INV 30-NOV-2002
DEFINITION Clona intestinalis cDNA, clone:citb012j08, full insert sequence.
ACCESSION AK116155
VERSION AK116155.1 GI:23588595
KEYWORDS FLI CDNA.
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Clona.
REFERENCE 1
AUTHORS Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T.,
Sasaki, A., Hamauchi, M., Awazu, S., Yagi, K., Sasakura, Y.,
Nakayama, A., Iehikawa, H., Inaba, K. and Satoh, N.
A cDNA resource from the basal chordate Clona intestinalis
Genes 33 (4), 153-154 (2002)
JOURNAL Genes 33 (4), 153-154 (2002)
MEDLINE 12203911
PUBMED 22191024
REFERENCE 2 (bases 1 to 1663)
AUTHORS Satou, Y. and Satoh, N.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail: satoh@esaid.cdn.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
Fax:81-75-705-1113)
COMMENT Clona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES
source Location/Qualifiers
1. 1663
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb012j08"
ORIGIN
Query Match 12.3%; Score 41.4; DB 3; Length 1663;
Best Local Similarity 46.4%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 156;

Qy 35 ACAACTCAATAGATGAACAGAGTAAGTGGGTAGTACTGTTTATAAAGAGAGTAAT 94
|||||
```

```
Db 1602 ATAAGTGTATTCAGAAAGCAGAACATATGTGTTCACAAAGCGTCAACAAATAAGGTAAT 1543
Qy 95 AAAGATACTATCATCATTTTGGAGCAATAAGGAGGAGAGATTCAGCAAAACAGTGTGCTT 154
1542 GACAGTTCATACCTTTTCTTGTAAAAAATATGACCCCTGCTTCATTTTACATTTGACCA 1483
Qy 155 ACAAGTGGAAAAACAAGTTAAACTAAAGTACCCCTCTCTTGAACAAGATCAATGCCACAG 214
1482 CGAATGGAATAGGTGTTAAATAGTAAACATTAATAATGTTACCCGAAATGACCCCTG 1423
Qy 215 TTGAGCTTTAGCCAGCCACATCATCATGTAATTTGCTTCTGATAGCCCTGTTTATAAA 274
1422 CTAAGCATCTGCTTATGCTGTAATTTCTGTGTTCTTGTGACCAACATAGTGTGAAC 1363
Qy 275 TTCTCTTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTCTACACAGAGAAA 325
1362 AGCTTCTGTAATCTCCACCACTTGCAATCAGTTCTGCTTCTGCTCTTAA 1312

RESULT 5
OSJN00175/c
LOCUS OSJN00175 99855 bp DNA linear PLN 10-FEB-2004
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0064D20,
complete sequence.
ACCESSION AL662973
VERSION AL662973.3 GI:32489560
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS Peng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,
Liu, X., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y.,
Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L. S., Yu, Z., Fan, D.,
Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J.,
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R.,
Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J.,
Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,
Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S.,
Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,
Li, J., Hong, G., Xue, Y. and Han, B.
Sequence and analysis of rice chromosome 4
Nature 420 (6913), 316-320 (2002)
JOURNAL Nature 420 (6913), 316-320 (2002)
MEDLINE 22337377
PUBMED 12447439
REFERENCE 2
AUTHORS Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q., Hu, X.,
Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Weng, Q. J., Zhang, L.,
Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y., Li, C.,
Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K., Zhou, B.,
Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P., Fu, G.,
Wang, S. X., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F., Tu, Y. F.,
Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X. Y.,
Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W., Wang, L. J.,
Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L., Zhu, F. H. and
Hong, G. F.
Direct Submission
JOURNAL Submitted (27-DEC-2001) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNBa0064D20.
COMMENT On Jul 9, 2003 this sequence version replaced gi:21912515.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
```

genscan (<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (<http://genemark.biology.gatech.edu/GeneMark/>), tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the complete sequence against NCBI none redundant protein database (nr) (<ftp://ncbi.nlm.nih.gov/blast/db>) and the EST database at NCGR.

FEATURES

source

1. 99855
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="4"
/clone="OSJNBa0064D20"
/clone_lib="CUGI-OSJNBa"
complement (3866. .6899)
/gene="OSJNBa0064D20.1"
complement (join(3866. .5543,5638. .6485,6573. .6899))
/gene="OSJNBa0064D20.1"
/codon_start=1
/protein_id="CAE04217.1"
/db_xref="GI:32489561"
/db_xref="UniProt/TREMBL:Q7X884"
/translation="MARGSAPLDGSLPPSRIVSEROAGLPRFEMPESAIGREIVSLG
EGCAPRYPGRSVFELPFSMAGLVPPFSSPFMDVLEPYDLQVAHLTPNAVMTLAIPTH
LCMEFTGPGVNLRIYPCVLRKMDWSDWFYTPDLAEARLRLSQPPAQASWVA
FDVQGVDAVLDRLAGLFGQLTGAMVYDLHRRRIAPLQRRAGAWETYSGEDFMR
THQGRWDVADEFKIVQVRLNLSVEASLI PQGVLPICSDPDRASILITIMQAVGS
ERAPAGHGSHRSGEOSTPGGRASGPRDGGPGSGPADARGKRKQEGTTPPSPP
RGGAARASRRSEGAFTPSOPEGERKKRLKMGCTSPRGNLISPPWNSFNRPSP
ETPSRRHSKSGQSAEDPATTEARRSDREADRLREAEAREARARQTEEA
QAEATGQAEATTSSEATRETDQEAARLREAEQAQVAVRAQEAARAEAREARL
RREAGTTSATDEAGASLGPTPTCDAQKPGPGDIPEGSTISGSPSAASSRRLL
PTFPVPLSAEPLQLALAVANTVLDLSAQVEALQAERAELDAARVAEEGRRSVEA
MVEVGHAKRHVSELEARKVLGEIAKEVEBERGAALIAITVMNEAOTLLQYGSW
EALGKLDAAARGVLDAARERAEATEVASRRREALEARMALEERACAVERDLA
DREAVAIREATLTAHEACAEESALRLREDALTERERALEAEAAQRLDASLRL
EAAREQAQRUEGAARAAALNORAELEARELDARERLDAARSGGAAGSDIARLAA
AEHTIADQLADSSAGEVALRAGEVGFPMFRDASRLDRAGRQAGLWGRRTAKYA
ANQGGLAQLSEITAGTLQRLPEELERTIKSSSRDLA"
complement (10918. .12816)
/gene="OSJNBa0064D20.2"
complement (join(10918. .11100,11189. .11344,11862. .11979,
12107. .12816))
/gene="OSJNBa0064D20.2"
/codon_start=1
/protein_id="CAE04218.2"
/db_xref="GI:38346571"
/db_xref="UniProt/TREMBL:Q7X902"
/translation="MGVFCSCSLIDLQGLWLIHPWRPRGCAVQTPSPPTDRLSAL
PDALLHTILSSUGRQMVQTVLSLKRWRHLRWSVCLDIDQREFAAASENWAISRDL
EKPEDFADNVLAYRCGSPAKLDTFLRICDRYHSLSRSDTDWRIRGLKSCSPREHLH
FVRYDSYLLEMHKLSNGYCLTKHLTVNLSLHCECFMEHITVTLLVLELRNCSLY
LQSI THPKLKNVLGHGPAMPSLVRASIRFSVVFCKCNFLVNLVNTLSLEISGFREM
VRLFNQVFEHAFKNLRTILLDRCPSPYNELLRHLLQNSPNLEKLTVCCKESKGL
ENRKSQHNQVNRKLNKSTETIYKOIDVDRELVDLLDVSGLHPKNTIALTKI"
complement (23800. .27420)
/gene="OSJNBa0064D20.3"
complement (join(23800. .24075,24708. .24800,24919. .24998,
25602. .25967,26575. .27420))
/gene="OSJNBa0064D20.3"
/codon_start=1
/protein_id="CAE04219.2"
/db_xref="GI:38346572"
/db_xref="UniProt/TREMBL:Q7X801"
/translation="MEPVAKRAFTGGGSAADVRLSSLPDALLHAVMSSLPARQWQ
TCVLSEKRWHLWSPSLNLDSEFLLPIYDRQKQWEDPTNLLPHEHAPTLDASIR
ADVAGKHGRVDNRWRCIGKYCPRVLDTIAVATDSRYRLPDLAGSCRLGRLHLSYV
ALHSGFARQVRQSPVRLCRLGHLRCLTFKSIESSTYLNRLVSDNSMGSGDSLAIASPR
LASRLVLAFLFHYTKNGSLNGANSLSVEASVAVKSGRTSPGEAMLLCGLFVSLSLEL
KGTQELLVFLVGLQIKQSWRLAFTLVEISFOAILHEKFSQSPFNRLTSLDNCF
QAEQADADVAALDSQWVRPILSLORLEDKVTFVQEPDELSMFYNLKRACCCSIASSA
NVVQKAWESHEHNDEHGRPSLAAPSQSPAAGLAAAPPVIVSTRGRERPLHGEARQR

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

CRRSLFSAGLSSPHVPAGRRRHICGDIVGEDAVAQ"
complement (29612. .32846)
/gene="OSJNBa0064D20.4"
complement (join(29612. .29827,29918. .30088,30550. .30590,
31036. .31096,31762. .32646))
/gene="OSJNBa0064D20.4"
/codon_start=1
/protein_id="CAE04220.2"
/db_xref="GI:38346573"
/db_xref="UniProt/TREMBL:Q7X877"
/translation="MDRLSALPDGLLHVMSPFAPARQMAQTCVLSKRWHLRWSVPSL
NLDIREFNLSKEEDDDDETGWKMKDFIANLLMFHAPTLDTFVCTGAVQAAVGN
RSRRNHRNDYDPAVADNRSDVGRWIRGGIKYCPRLDVAVAPSGSGCRPPDLGS
GSCFRLRLRSLCSFALDSGFARQVRDSCPVLRCELHRCHIEFHSISSTLUNKLVIE
GCIGCSLAIASAPRLASLCLDLSYGAYKGVSLNTISLVEASVTLVNVPQISPEGEAM
LLCGLFNTVNLLEAIAHAKLGGVLGKLSAVGAKGGRERRKFTGCCNVNKKVLIHE
KDFEFSFNRLTSLSDCFQMGDLKEKFKALGRLLKCPNLKLEKLTQHCWFLSGST
QRAEWETTSGLCLQSQDQVTFHCQKLKLIETKHDSLCCDHLQLQLMWGFWKDLKA
TILLTKI"
complement (37808. .41919)
/gene="OSJNBa0064D20.5"
complement (join(37808. .37912,38701. .38871,38985. .39135,
41339. .41919))
/gene="OSJNBa0064D20.5"
/codon_start=1
/protein_id="CAE04221.2"
/db_xref="GI:38346574"
/db_xref="UniProt/TREMBL:Q7X6F3"
/translation="MGPRRRAPATVLEINVVTPSPPGHYDGYTMPDLIITSRAWRR
LTRHLSHAWLDGGFGEOGDCPLLEDALRLRCAMAPGRIRICGSLRTLVLHYTGC
GGDAGGDEEEETIYISAPRLASVRKJTSYACRHGVSFDSSTADSLVEASIRVG
RRRRALPTGVAEVLLAGVNVVTTLSKVTRQRTGTVSLGVNSLVEASIDVNSQOM
SPSGAMILGALFSAINTLELGRAMAILDEELDKFPLFNRLNLSLHCCURDKNLS
DRFKALGRLLQKSLNLEKLTLDQDFWASWCRNTSSVLGKIQVLLVLLFTPIILLVASDL
N"
complement (45448. .46707)
/gene="OSJNBa0064D20.6"
complement (join(45448. .45516,45762. .46130,46270. .46707))
/gene="OSJNBa0064D20.6"
/codon_start=1
/protein_id="CAE04222.2"
/db_xref="GI:38346575"
/db_xref="UniProt/TREMBL:Q7X5X6"
/translation="MADRTTATMI PAASSSRKRARVTTGAGDGGGGGEGGRLGELPD
ELLLSILSCLTTRQAVQTVLSLRWRHLWRSTPREDVDLAEFARPPSPAPWLLHGRG
STDPEWRLRGFTFARLLMTHAAPVLDAPFLRVATPFRRADVESWAFAGDLLSGCPAL
VDMELDRCKCFHELSSATLSRLAMESCLWMRPPSGTNGDRTVSVVAPRLAYLRLLTF
GHGCKVFRFSGDSISEVSRINGFNILNLPRLRMPNVVTLRLSGFGPTSVSARMF
RKFRPPPSDNLALG"
47424. .50668
/gene="OSJNBa0064D20.7"
Join (47424. .47697,48191. .48465,48691. .48855,48923. .49012,
50294. .50668)
/gene="OSJNBa0064D20.7"
/codon_start=1
/protein_id="CAE04223.2"
/db_xref="GI:38346576"
/db_xref="UniProt/TREMBL:Q7X8C0"
/translation="MGCGGGENRVNDHDDDDDLWELAAALADVTLLVGVKVGSGKS
ATANSILGDEAPESKCSYAGVTQTCQKSTTVQDGLRTINVIDTGPLFDMIDKAEAD
VRVYVKCMDAKDGIIHAMLVFSATSRFSCEDEKTIETLKSFFGDKILDHMLVYTR
GDEVYKSNMLSDSAFTYLDILKLFENRVLPENKTSSTQDRQQRKQKMLDAVD
FLVSNHGKFPNSQLFTQIQEVHHRQKDSANSEVSMOETDSYISLITKMWBEKLNFT
ILVMEQQLLKQEARLDIQENMTKALRSEEDIRRLRSLSEKAEISNWNARENKRFR
ESEKASKEQEKQTEAIIQKLKEMKEKREEREIRRLRDDLEKEREERQKQSGCIIIL
"
complement (51830. .56465)
/gene="OSJNBa0064D20.8"
complement (join(51830. .52739,52843. .52970,54357. .54818,
56012. .56229,56348. .56465))
/gene="OSJNBa0064D20.8"
/codon_start=1
/protein_id="CAE04224.2"

```
/db_xref="GI:38346577"
/db_xref="UniProt/TREMBL:Q7X729"

Query Match
Best Local Similarity 51.3%; Score 41.4; DB 8; Length 99855;
Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 37 AACTCAATAGATGAACACAGATGAATGGGTAGTCTGTTTATRAAGAGAGTAATAA 96
Db 83615 AAATTAGTACCAACACACATAGTCATGCTCTTTCTTCATGAATGAATAATGTCGTGT 83556

QY 97 AGATCTATCATCTTTGAGCAATAGGAGGAGAGATTCAGCAACAGTGTGCTTAC 156
Db 83555 ATAATATCTGAATTTGTGATAATATTAAAGATGTAACAGTAATGAATGCTGTTA 83496

QY 157 AAGTGGAAAAACAAGTTAACTAAAGTACCCCTCTCTGACAGATCAATGCCACAGTT 216
Db 83495 AAGTGGAAAAGTTGTTCTACAGATATATCTCTCATCTCCCTAAATGTTGACACCGTT 83436

QY 217 GAGCTTT 223
Db 83435 GACTTTT 83429

RESULT 6
OSJN00292/c
LOCUS
DEFINITION
Oryza sativa genomic DNA, chromosome 4, BAC clone: P0076017,
complete sequence.
ACCESSION
BX548156
VERSION
BX548156.1
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X.,
Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Wang, Q.J., Zhang, L.,
Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C.,
Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B.,
Chen, Z.H., Hao, P.S., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G.,
Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F.,
Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y.,
Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., and
Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and
Hong, G.F.
Direct Submission
Submitted (08-JUL-2003) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: P0076017.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Egenes (http://www.softberry.com/),
genscan (http://CCR-081.mit.edu/GENSCAN.html), Genemark-ES (
http://genemark.biology.gatech.edu/Genemark/), TRNscan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/TRNscan-SE/), searches of the
complete sequence against NCBI non-redundant protein database (nr)
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES
Location/Qualifiers
1..121130
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
source
```

```
/db_xref="taxon:39947"
/chromosome="4"
/clone="P0076017"
/clone_lib="RGP-OSJNP"
/complement(4269..7724)
/gene="P0076017.1"
/complement(join(4269..4451,4540..4695,5213..5330,
5458..6125,7695..7724))
/gene="P0076017.1"
/codon_start=1
/protein_id="CAE02503.1"
/db_xref="GI:32479668"
/db_xref="UniProt/TREMBL:Q7X902"
/translation="MASSFRKRTKWLRIILHPMPRGVAVQTPSPATRLSALPDAL
LHTILSLKGMVOTSVLSKRWHLWRSVPCLDIDREFAAASENAI,SRSDLEKFE
PDADNVLYRCGSPAKLDTFLRICDHYSLRSSDTRMIRGLKCSREFPHLHFDYR
YSYLLEMKHGSYNGCLTKHITNVSLECFMEHITVCTLEVLNRCSTLYLQRI
THPKLNLVHGPAVMPSLRASIRFVPCPKNFNLNFTVLELSGFRFMRLE
FNQVEFHAFKNLRLTLDDCRPSYNNELRHLLQNSPNLEKLTVCCHCKFSKSLWRK
SSQHKQNVNCRKLKSTETIYKDIIDVRELVDLLDVSGHLPKNTIALTKI"
11169..13339
/gene="P0076017.2"
/join(11169..11467,12805..13339)
/gene="P0076017.2"
/codon_start=1
/protein_id="CAE02504.1"
/db_xref="GI:32479669"
/db_xref="UniProt/TREMBL:Q7X801"
/translation="WAPPIKSLSQGRPIAPVWSHLLYARMFPQBSVLICEYGVPH
RNPRIASFHNPFFYQTHRRHMSGFGFNSFPFKFTNNGVCGIFGIRQARIPRI
GWSYTEGRRRAATTFGWRQRAYGEQRRSRQLATGRLLIALLPRAKERRSSEG
EREMHYRAETDVVTSGRRGSELGLGPGKGTAPGVPHOLLARTPASPARAVEGG
EAGGSKAAKARWNRGNEREGDGAARNRQCRFREKEPTLAVKESELAVIDACSQGD
R"
complement(16366..20771)
/gene="P0076017.3"
/complement(join(16366..16494,18059..18151,18270..18367,
18906..18969,19067..19237,19926..20771))
/gene="P0076017.3"
/codon_start=1
/protein_id="CAE02505.1"
/db_xref="GI:32479670"
/db_xref="UniProt/TREMBL:Q7X877"
/translation="MEPVAKRARTGGGSAADVRLSSLPDALLHAVMSSLPARQMVQ
TCVLSKRWHLWRSVPSLNDREFLLPIYDRWQKMEFTTNLLMFHAPTLDAFSIR
ADVAGKCHGVDRMIRCGIKYCPVRLDIATVDSRYRLPDLASCSRLGRHLISYV
ALHSGFARQDSCPVLCLELHRLKTFSHIESSTLNRLVIEDSGGSDSLAISAPR
LASLRLVAPLFTHYKNGVSLGANSLSVEASVAVKSGRTSPEGEAMLLGLFSVSSLEL
KGIOBLAILHEKFDKSGSFHNLRTSLCLENSETLGDADDAALDQWQVRPILSLQRLD
LQHCWFLGSGTORAKRKTSTGLCLENSETLGDADDAALDQWQVRPILSLQRLD
KYTFVQEPDELSMFYNLKRACCCSIASSANVISERLKRGEGETDLDSDVPDPHDFNT
SQCPNLKVIKIKLI"
complement(22963..25997)
/gene="P0076017.4"
/complement(join(22963..23178,23269..23439,23901..23941,
24387..24447,25113..25997))
/gene="P0076017.4"
/codon_start=1
/protein_id="CAE02506.1"
/db_xref="GI:32479671"
/db_xref="UniProt/TREMBL:Q7X6F3"
/translation="MDRLSALPDGLLHVMSPFAPARMAOTCVLSKRWHLWRSVPSL
NLDIREFNLSKEEDDDDETWGKKDFIANTLMFHAPTLDAFSIR
RSRRNRNDVDYPAVANRSDRVGRWIRGGIKYCPVRLDVAVASGSCRRPPDGLS
GSCFRLRLRSLSCFALDSGFARQDSCPVLCLELHRLKTFSHIESSTLNKLVIE
GGIGSLASLAPLASLCLDSYGKNGVSLNTISLVEASTINLVNFQISPEGAM
LLCGIFNVNLTLEIAHAKLGGVLGKLSAVGSAKGGRRKRFYGCNNKVKILHE
KDFKFSFNLTSLDLCFCQGMGLKKEFKALGRLEKCPNLKLTLOHCWFLSGST
QRAEMETSTGLCQSQQVTFHCQKLIKLEIKHDSLECCDHLQLQLMGWFWKDLKA
TIIITKI"
complement(31159..35270)
/gene="P0076017.5"
/complement(join(31159..31263,32052..32222,32336..32486,
```



```
34690..35270))
/gene="P0076017.5"
/codon_start=1
/protein_id="CAE02507.1"
/db_xref="GI:32479672"
/db_xref="UniProt/TREMBL:O7X5X6"
/translation="MGPRRRAPATVLEINVTPSPYPPGHYDGYTMPDLIIITRAWR
LTRLHLHAWLDGFGQEGDGCPLLEDIALRCACWAPFRIRGCSLRLVLHYTGC
GGDAGGDEEEETLVISAPRLASVRKITSYACRHGVSGDGLSVASIRVG
RRRRRLPTGEVAAELAGLVNVTLSKVYRQDGTSLVGNLSVEASIDVSCNM
SPSGEAMLCALFSATNLKGTGRAMILDELDKPLFNLRNLHLCHCLDKGNLS
DRFKALGRLLQKLSNLEKLTLDQFWSWCRNTSSVLGKIQVLVLILFIILLHVASDL
N"
complement(38799..40058)
/gene="P0076017.6"
/complement(join(38799..38867,39113..39481,39621..40058))
/gene="P0076017.6"
/codon_start=1
/protein_id="CAE02508.1"
/db_xref="GI:32479673"
/db_xref="UniProt/TREMBL:O7X8C0"
/translation="WADRTTATMIPASSSRKRARVPTTGADGGGGEGRIGELPD
ELLSTLSCLTTRQAVOTVSLGRWRHLWRSTPRFDVLAERFPPSPAPMLHGRG
STDPERLKGFTALILMVAAPVLDAFLRLRVATPHRRADVESWSAFAGDLISGCPAL
VMDLDRCKCFPHELSSATLSLAMESCLWRRPSTGNGDRVTSVAPRLAYLRLLTP
GHDCCKVRFSGDSLSISVSIKGGFNLIILFLRLMRMPNVTILRLSGFGPTSVSARMF
RKPRPPQSDNLALG"
40499..44019
/gene="P0076017.7"
/join(40499..41048,41542..41816,42042..42206,42274..42363,
43645..44019)
/gene="P0076017.7"
/codon_start=1
/protein_id="CAE02509.1"
/db_xref="GI:32479674"
/db_xref="UniProt/TREMBL:O7X7Z9"
/translation="MEEVWLDLLEIMGYTKLARTGEAHHAANLFAVSDTSLSSFR
GKLSAACINHLHSPSSSLPPLFPVSUSATPVISQLSIDPIMGGGGENRV
DHKDDDDDEGLAALADVTLVLKVGSGKSATANSILGDEAFSPKSCSVAGVTQT
COKSTQDQGLARTINVIDTPLGDMDKAEDVREIVKCMADKAGIHAIVTFS
ATSRFCEKTEITLKSFPFKILDKMLILVTRGDEVGETSWKNSLSDSAPTLYQD
ILKLPNRVLFNKTSOTDQAOAKMLDAVDVFWSSNHGKPPSNOLFOTOEVHH
RUKDANSVYSNQETDSYISLITKMEVKLANGTILMEQQLKEQEARLDLONETK
AIRSEDRRLRLSLEKASQNSNARENKRPRESEKASKEQKQTEABIQKLKEM
EKDREEREIRLRDLDEKEREERQKSGCII"
complement(45181..49816)
/gene="P0076017.8"
/complement(join(45181..46090,46194..46321,47708..48169,
48735..48764,49363..49580,49699..49816))
/gene="P0076017.8"
/codon_start=1
/protein_id="CAE02510.1"
/db_xref="GI:32479675"
/db_xref="UniProt/TREMBL:O7X619"
/translation="MDVESCASSPPDAVDWRGCPCEPRBHGWRRAAVFLGIQAF
ETMALAAGNLTIVFGEMHPLSQANVNTNFGVTFILLSILGFLSDSLGCFWT
MLIFGVELSVFNDTATATGTLISVQAHLPQLKPPCNMAATDGCCEQARIGKASI
FFAALYLVALGSGCLKNMIAHGADQFAAAGGGAADNAKRLSTYFNSYFSCA
GELVALTALVQTHSGMDVFGFISAAAGLISVLSGAAFRNKPPOGSIPTFIAR
VFVAATKQKQICPSSSDPVNAGCEPAHLAGGSPRHASKERFLDKACIRAEQGN
TPKESPMRLCTAEVRSQAKTLAVAPIFACTIVENTVLAQLQTFESVQOQSMDALGQ
AGSFRIPASLOAIPYAMLLALVPAYELLVPLMRATGASGITPLQRIUGLCTV
PLSWAAATVHRRRLDLSAGAPPRAWSLVIVQFLVFGVSMFTVAGLIEFFKY
QARGAGQFLLTALCYSAFGFLSSVLVSNRVNTRSGGGHGGHGLGNDLIDKDR
LDLFYMWLAVLVNFFCYLLCARWYNSGGADGDGDASASQAAVEEDGNGKEII"
complement(64136..69177)
/gene="P0076017.9"
/complement(join(64136..64335,64429..64596,64680..64899,
65493..65683,65791..65884,66179..66422,67236..67333,
67431..67522,67617..67731,67812..67958,68953..69177))
/gene="P0076017.9"
/codon_start=1
/protein_id="CAE02511.1"
```

```
/db_xref="GI:32479676"
/db_xref="UniProt/TREMBL:O7X914"
/translation="MSQPAVAAGAAGAAQQPQRAQAARQAGGGLGQTIAGIVRMAVF
WYFAAKFGPKRPPAPGPMGLMNLFGKPELDMWMLYSENEKFNDFSNEDALVWHEAN
IPYAVAGPTSTRTHLTLYVPSAEIKNGSLVAHVYFARSQYPDVTDPEYEQSSAFGR
THPVAVYLKPKDGKKKSLIGSSSENEQPPPKENKNSVDKDEGVVEIYSYKPNVTI
NLVDDPTNTDLNDVPSSNNYPTVTFNFEFLRLDLALNETVELPLNLEVGLISM

Query Match      12.3%; Score 41.4; DB 8; Length 121130;
Best Local Similarity 51.3%; Pred. No. 2.8;
Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY      37  AACTCAATAAGATCAACAGATGAATGGGTTAGTGACTGTTTATAAAGAAGAGTAATAA 96
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      76966  AAATTAGTAACCAACACACTAGTCATGCTCTTTCTTATAATGAATAATGTGGTGT 76907
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      97  AGATACTATCATCTTTGAGGCAATAAGGAGGAGAGATTTCAGCAACAGTGTGCTTAC 156
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      76906  ATAAATTATCTGAATTTGTGATAATATTAGAANTTAACAGATAGATTAACTGCTTTA 76847
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      157  AAGTCGAAAAAAGTTAACTAAAGTGACCCCTCTTTGACAAAGATCAATCCACAGTT 216
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      76846  AAGTTGAAAAAGTTGTTCTTACAAAGATATATCTCCTCATTCCTAAATGTTTGACACCGTT 76787
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      217  GAGCTTT 223
DB      76786  GACTTTT 76780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AL928605  Mouse DNA sequence from clone RP23-139P14 on chromosome 4, complete
LOCUS     sequence.
DEFINITION
ACCESSION AL928605.22 GI:51965306
VERSION    AL928605
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 209024)
AUTHORS    Tracey A.
TITLE       Direct Submission
JOURNAL     Submitted (08-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
COMMENT     Cambridgehire, CB10 1SA, UK. E-mail enquiries:
             humquery@sanger.ac.uk
             On Sep 9, 2004 this sequence version replaced gi:50949682.
             Sequence from the Mouse Genome Sequencing Consortium whole genome
             shotgun may have been used to confirm this sequence. Sequence data
             from the whole genome shotgun alone has only been used where it has
             a phred quality of at least 30.
             ----- Genome Center
             Center: Wellcome Trust Sanger Institute
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: humquery@sanger.ac.uk
             -----
             During sequence assembly data is compared from overlapping clones.
             Where differences are found these are annotated as variations
             together with a note of the overlapping clone name. Note that the
             variation annotation may not be found in the sequence submission
             corresponding to the overlapping clone, as we submit sequences with
             only a small overlap as described above.
             This sequence was finished as follows unless otherwise noted: all
             regions were either double-stranded or sequenced with an alternate
             chemistry or covered by high quality data (i.e., phred quality >=
             30); an attempt was made to resolve all sequencing problems, such
             as compressions and repeats; all regions were covered by at least
             one plasmid subclone or more than one M13 subclone; and the
             assembly was confirmed by restriction digest, except on the rare
             occasion of the clone being a YAC.
             The following abbreviations are used to associate primary accession
```


numbers given in the feature table with their source databases: Em., EMBL, Sw., SWISSPROT, Tr., TREMBL, Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-139P14 is from the RPCI-23 Mouse BAC Library. constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES
source
Location/Qualifiers
1. .209024
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-139P14"
/clone_lib="RPCI-23"

ORIGIN
Query Match 12.1%; Score 40.6; DB 10; Length 209024;
Best Local Similarity 53.5%; Pred. No. 4.3;
Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 22 CCGGTTTTTACTCACTCAATTAAGATGAACAGAAATGAATGGTTAGTCACTGTTTAT 81
Db 192057 CTGATTATTTTCATCAAGGAAGAAAGAAACAGAGGATTAAGGAGAAATCAGAAAGAA 192116

QY 82 AAGAAGAGTAATAAGATATCTATCATCTTTGAGCAATAAGAGGAGGAGAGATTCCAGC 141
Db 192117 AAGAAGATGAAGGAAAGAAAGAAAGAAAGGAGTAAGGAGGAGGAGGAGGAGG 192176

QY 142 AACAGTGTGTTACAGAGTGGAACAAAGTTAACTAAA 180
Db 192177 TAGGAGTGGGGGAAAGAAAGAAAGTTAAAGAAAGAAAGAA 192215

RESULT 8
AL583823 3/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL583823 Accession AL583823
Fragment Name Begin End
AL583823_0 1 110000
AL583823_1 100001 210000
AL583823_2 200001 310000
AL583823_3 300001 360756
Continuation (4 of 4) of AL583823 from base 300001 (AL583823 Homo sapiens chromosome 9 contig 3)

Query Match 12.0%; Score 40.4; DB 2; Length 60756;
Best Local Similarity 55.8%; Pred. No. 5.3;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 197 ACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCATGTAAATTCGTTTCCT 256
Db 33572 AAAAGATGATGCCCTCACTAGTTTACACAGATGATTTAGCATGATCAAGATTAACCT 33513

QY 257 GATAAGCTCTTCATAAATTCCTTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTTAC 316
Db 33512 GGTTAGCTTTTAAATTCCTTTCCAGTCTTACTTCTAAAGAAAGTCACTGTTT 33453

QY 317 AGACAGAAAGATGCCATT 334
Db 33452 AGAAGAAAGAGATATT 33435

RESULT 9
AL354710
LOCUS
DEFINITION
Human DNA sequence from clone RP11-65N13 on chromosome 9, complete sequence.
ACCESSION
AL354710
VERSION
AL354710.17
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 131708)
Lloyd, D.

Direct Submission
Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Feb 1, 2002 this sequence version replaced gi:18375790.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

RP11-65N13 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-65N13. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RP11-65N13 is at 131708 in this sequence. The true left end of clone RP11-184822 is at 91097 in this sequence. The true right end of clone RP11-366020 is at 2000 in this sequence.

FEATURES
source
Location/Qualifiers
1. .131708
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-65N13"
/clone_lib="RPCI-11.1"
68692..68732
/note="Sequence from overlapping clone RP11-180F6 (AL583823). Assembly confirmed by restriction digest."
69053..69060
/note="Sequence from overlapping clone RP11-180F6 (AL583823) and RP11-233N7 (AL445234). Assembly confirmed by restriction digest."
80432..81341
/note="Sequence from overlapping clone AC026155 sequenced by BCM and RP11-180F6 (AL583823). Assembly confirmed by restriction digest."
89032..89102
/note="Sequence from overlapping clone RP11-180F6 (AL583823). Assembly confirmed by restriction digest."

misc_feature
/note="Sequence from overlapping clone RP11-180F6 (AL583823). Assembly confirmed by restriction digest."
69053..69060

misc_feature
/note="Sequence from overlapping clone RP11-180F6 (AL583823). Assembly confirmed by restriction digest."
80432..81341

misc_feature
/note="Sequence from overlapping clone RP11-180F6 (AL583823). Assembly confirmed by restriction digest."
89032..89102

ORIGIN
Query Match 12.0%; Score 40.4; DB 9; Length 131708;
Best Local Similarity 55.8%; Pred. No. 5;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 197 ACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCATGTAAATTCGTTTCCT 256
Db 114544 AAAAGATGATGCCCTCACTAGTTTACACAGTATTTAGCATGCATGAAGAATTACCT 114603

```
QY 257 GATAAGCCTGTTTCATAAATCTCTTTGCAAGCTCTGCTACTTACCAGAGTCGTGCTAC 316
Db 114604 GGTAGCTTTTAAAAATTCATTTCCCAAGTCTTACTTCTTAAAAAAGAAGTCAGTTT 114663

QY 317 AGACAGAAAGATGGCATT 334
Db 114664 AGAAGAAAAGAGTATT 114681

RESULT 10
AC026155/c
LOCUS
DEFINITION Homo sapiens chromosome 12 clone RP11-180F6, WORKING DRAFT
AC026155
VERSION AC026155.23 GI:14547355
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muzny,D.M., Adams,C., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunga,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimarge,K., Blankenburg,K., Bomin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burke,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neale,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Peters,R., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,F., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y.F., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 174294)
Worley,K.C.
Direct Submission
Submitted (21-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:13811703.
----- Genome Center
Center: Baylor College of Medicine
```

```
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HALP
Center clone name: RP11-180F6
----- Summary Statistics
Sequencing vector: Plasmid; M77789
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 176003 bases at least Q40
Consensus quality: 179810 bases at least Q30
Consensus quality: 181603 bases at least Q20
Estimated insert size: 178132; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 8.6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 35024: contig of 35024 bp in length
* 35025 35124: gap of unknown length
* 35125 56161: contig of 21037 bp in length
* 56162 56261: gap of unknown length
* 56262 77352: contig of 21091 bp in length
* 77353 77452: gap of unknown length
* 77453 103958: contig of 26506 bp in length
* 103959 104058: gap of unknown length
* 104059 119854: contig of 15896 bp in length
* 119855 120054: gap of unknown length
* 120055 130550: contig of 10496 bp in length
* 130551 130651: contig of 12053 bp in length
* 130652 142703: contig of 12053 bp in length
* 142704 142803: gap of unknown length
* 142804 149562: contig of 6759 bp in length
* 149563 149862: gap of unknown length
* 149863 157532: contig of 7870 bp in length
* 157533 157632: gap of unknown length
* 157633 168762: contig of 11130 bp in length
* 168763 168862: gap of unknown length
* 168863 171961: contig of 3099 bp in length
* 171962 172061: gap of unknown length
* 172062 174294: contig of 2233 bp in length.
FEATURES
Location/Qualifiers
1..174294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-180F6"
ORIGIN
Query Match 12.0%; Score 40.4; DB 2; Length 174294;
Best Local Similarity 55.8%; Pred. No. 4.9;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 197 ACAAGATCAATGCCAGTTGAGCTTTAGCCAGCCACATCATGTAATTCCTTCCT 256
Db 117449 AAAAGATGATGCCCTCACTAGTGTTCACACAGTGTATTTAGCATCATGAATTACCT 117390
QY 257 GATAAGCCTGTTTCATAAATCTCTTTGCAAGCTCTGCTACTTACCAGAGTCGTGCTAC 316
Db 117389 GGTAGCTTTTAAAAATTCATTTCCCAAGTCTTACTTCTTAAAAAAGAAGTCAGTTT 117330
QY 317 AGACAGAAAGATGGCATT 334
```

Db 117329 AGAAGAAAAGAGATATT 117312
|||||
RESULT 11
AE014175 2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AE014175 Accession AE014175
Fragment Name Begin End
AE014175_0 1 110000
AE014175_1 100001 210000
AE014175_2 200001 310000
AE014175_3 300001 404829
Continuation (3 of 4) of AE014175 from base 200001 (AE014175 Mus. musculus piebald deletion)
Query Match 12.0%; Score 40.2; DB 10; Length 110000;
Best Local Similarity 48.9%; Pred. No. 5.7; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 113
Db 15331 TTTTATAGCTAAATTTAGAAAATTCCTTCATCTCTGGTATGTCTATTATAAATTAT 15272
|||||
Qy 26 TTTTACTACAATCAATGAATGAACAGAAATGAATGGTTAGTGACTGTTTATAAAG 85
Db 15331 TTTTATAGCTAAATTTAGAAAATTCCTTCATCTCTGGTATGTCTATTATAAATTAT 15272
|||||
Qy 86 AAGAGTAATAAGTACTATCATCTTTGAGGCAATGAAGGAGGAGAGATTGAGCAAC 145
Db 15271 CAGACATGTTTCTTCTTCAGTCTAGAGGTGAGGTTATTGAGAGGGAGAGACCACAGTG 15212
|||||
Qy 146 AGTGTGCTTCAAGTGAAGAAACAGTTAAACTAAAGTGACCCCTCTTGCAGCAAGTCA 205
Db 15211 AGTGTGAGAACAGTGCAGACCAACATGACCCCTGGGCAAGGCTTTCATTATAAGAACA 15152
|||||
Qy 206 ATGCCAGTTGAGCTTTAGCCAGCCACATCATCATGATAA 246
Db 15151 GGGAGAGGAGGAGTGAAGGCGTGCCACATCATTCACA 15111
|||||
RESULT 12
AC114410/c
LOCUS
DEFINITION Mus musculus, clone RP23-151K8, complete sequence.
ACCESSION AC114410
VERSION AC114410.9 GI:34536773
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1. (bases 1 to 201783)
Birren, B., Nusbaum, C. and Lander, E.
Unpublished
2. (bases 1 to 201783)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galsgan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3. (bases 1 to 201783)
Birren, B., Nusbaum, C. and Lander, E.
Direct Submission
Submitted (15-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4. (bases 1 to 201783)
Birren, B., Nusbaum, C. and Lander, E.
Direct Submission
Submitted (09-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2003 this sequence version replaced gi:33667216.
All repeats were identified using RepeatMasker:
Smith, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23439
Center clone name: 151_K_8
----- Location/Qualifiers
1. .201783
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-151K8"
/clone_lib="RPCI-23 Female Mouse BAC"
1. .7101
/note="wgs_end extension"
clone end:SP6
complement(90..201)
/rpt_family="PB1D10"
443..539
/rpt_family="B1P"
546..574
/rpt_family="B1P"
/rpt_family="(TCCA)n"
complement(686..802)
/rpt_family="PB1D9"
1155..1211
/rpt_family="(TA)n"
1436..1618
/rpt_family="B2_Mm2"
2164..2185
/rpt_family="AT rich"
complement(3339..3466)
/rpt_family="B1P"
complement(4368..4448)
/rpt_family="L1_MM"
7102..7107
/note="clone boundary"
clone end:SP6
site:EcoRI
9053..9152
/rpt_family="(TA)n"
12029..12129
/rpt_family="MTE"
12236..12345
/rpt_family="PB1D7"
12752..12887
/rpt_family="(TC)n"
complement(13407..13609)
/rpt_family="B3"
14127..14167
/rpt_family="(TA)n"
14167..14193
/rpt_family="(CA)n"

```
repeat_region 14235..14272 /rpt_family="(CA)n"
repeat_region 14521..14558 /rpt_family="(TTA)n"
repeat_region 14935..14963 /rpt_family="AT_rich"
repeat_region 15589..15629 /rpt_family="(TTTG)n"
repeat_region complement(15632..15765) /rpt_family="B1_MM"
repeat_region 15919..16063 /rpt_family="B1_MM"
repeat_region 16068..16090 /rpt_family="B1_MM"
repeat_region 16143..16166 /rpt_family="(CAAC)n"
repeat_region 16177..16199 /rpt_family="(CAAA)n"
repeat_region complement(17349..17514) /rpt_family="RSINE1"
repeat_region 18604..18631 /rpt_family="(A)n"
repeat_region 18696..18735 /rpt_family="(A)n"
repeat_region 19686..19753 /rpt_family="CT-rich"
repeat_region 22500..22598 /rpt_family="(CT-rich)"
repeat_region 22891..22945 /rpt_family="(TTTC)n"
repeat_region 23058..23105 /rpt_family="CT-rich"
repeat_region complement(24470..24595) /rpt_family="RSINE1"
repeat_region 26176..26465 /rpt_family="(TATA)n"
repeat_region 26473..26541 /rpt_family="B4A"
repeat_region 26917..26939 /rpt_family="(CA)n"
repeat_region 28766..28877 /rpt_family="(CAAAA)n"
repeat_region 28878..28909 /rpt_family="PB1D9"
repeat_region 29243..29455 /rpt_family="(CAAA)n"
repeat_region 30293..30336 /rpt_family="B4"
repeat_region complement(30557..31009) /rpt_family="AT_rich"
repeat_region 31219..31213 /rpt_family="L1"
repeat_region complement(31159..31213) /rpt_family="L1"
repeat_region 31219..32472 /rpt_family="L1"
repeat_region 31219..32472 /rpt_family="Lx4"
repeat_region complement(33069..33221) /rpt_family="L1_MM"
repeat_region 33438..33470 /rpt_family="(CAAA)n"
repeat_region complement(36123..36255) /rpt_family="B1F"
repeat_region 37393..37422 /rpt_family="AT_rich"
repeat_region 40393..40422 /rpt_family="AT_rich"
repeat_region 40692..40712 /rpt_family="(TA)n"
repeat_region 41263..41297 /rpt_family="AT_rich"
repeat_region 42833..42861 /rpt_family="AT_rich"
repeat_region 43372..43402 /rpt_family="AT_rich"
repeat_region complement(44537..44680) /rpt_family="AT_rich"
repeat_region 44680..44711 /rpt_family="B1_MM"
repeat_region complement(44781..44978) /rpt_family="B1_MM"

/rpt_family="URRIA"
44993..45026 /rpt_family="AT_rich"
45864..45939 /rpt_family="CT-rich"
45948..45982 /rpt_family="(CA)n"
complement(47163..47222) /rpt_family="5S"
47927..48187 /rpt_family="Lx8"
48211..48294 /rpt_family="Lx8"
48808..48990 /rpt_family="B3"
complement(50184..50368) /rpt_family="B3"
50559..50602 /rpt_family="(TAA)n"

Query Match 12.0%; Score 40.2; DB 10; Length 201783;
Best Local Similarity 48.9%; Pred. No. 5.4;
Matches 108; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 26 TTTTCTACTCAACTCAATAAGATGAACAGATGAATGGGTAGCTGTTTATAAG 85
DB 87178 TTTTATAGCTAAATTTAGAAAATTTGCTTCAATCTCTGGTATTCTTATAAATTAT 87119
QY 86 AAGAGTAAATAAGATATCTATCATCTATTGAGGCAATAAGGAGGAGAGATTCAGCAAC 145
DB 87118 CAGACATGTTTCTTCTTCTAGCTCTAGAGGTGAGGTATTGAGAGGAGAGAGACACAGTG 87059
QY 146 AGTGTGCTTACAAGTGGAAACAAAGTTAAACTAAAGTACCCCTCTCTGCAAGATCA 205
DB 87058 AGTGTGAGAAACAAGTGCAGCACCATGACCTGGGCAAGGCTTTCCATTATAAGACA 86999
QY 206 ATGCCACAGTTGAGCTTTAGCCAGCCACATCATCATGTTAAA 246
DB 86998 GGGAGAGGGAGGAGTGAAGGCGGTGCCAACATCATTCACA 86958

RESULT 13
CR318618/c
LOCUS
DEFINITION
DANIO rerio clone CH211-3917, *** SEQUENCING IN PROGRESS ***, 10
unordered pieces.
ACCESSION
CR318618
VERSION
CR318618.1 GI:44864762
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 177655)
McLay, K.
Direct Submission
Submitted (27-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
***** Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
***** Project information
Center project name: zC3917
***** Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 175828 bases at least Q40
Consensus quality: 176161 bases at least Q30
Consensus quality: 176379 bases at least Q20
```

Insert size: 176755; sum-of-contigs
 Insert size: 188044; 8.4% error; agarose-fp
 Quality coverage: 9.57x in Q20 bases; sum-of-contigs Quality
 coverage: 9.27x in Q20 bases; agarose-fp

***** NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 57111: contig of 57111 bp in length
 57112 gap of 100 bp
 57212 contig of 15863 bp in length
 73075 gap of 100 bp
 73175 gap of 7082 bp in length
 80257 gap of 100 bp
 80357 gap of 100 bp
 89661 contig of 9304 bp in length
 89761 gap of 100 bp
 98899 contig of 9139 bp in length
 98900 gap of 100 bp
 99000 contig of 6066 bp in length
 105066 gap of 100 bp
 105166 contig of 25086 bp in length
 130252 gap of 100 bp
 130352 contig of 36588 bp in length
 166940 gap of 100 bp
 167040 contig of 5095 bp in length
 172135 gap of 100 bp
 172235 contig of 5421 bp in length.

FEATURES

source

1..177655 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="CH211-3917"
 /clone_lib="CHORI-211"
 misc_feature 1..57111
 /note="assembly fragment:01750
 fragment_chain:1
 clone end:SP6
 vector side:left"
 misc_feature 57212..73074
 /note="assembly fragment:00526
 fragment_chain:1"
 misc_feature 73175..80256
 /note="assembly fragment:00319
 fragment_chain:1"
 misc_feature 80357..89660
 /note="assembly fragment:00418
 fragment_chain:2"
 misc_feature 89761..98899
 /note="assembly fragment:00153
 fragment_chain:2"
 misc_feature 99000..105065
 /note="assembly fragment:00044
 fragment_chain:2"
 misc_feature 105166..130251
 /note="assembly fragment:00847
 fragment_chain:2"
 misc_feature 130352..166939
 /note="assembly fragment:01200
 fragment_chain:2"
 misc_feature 167040..172134
 /note="assembly fragment:00235
 fragment_chain:2"
 misc_feature 172235..177655
 /note="assembly fragment:00087
 fragment_chain:2"

ORIGIN

Query Match 11.8%; Score 39.8; DB 2; Length 177655;
 Best Local Similarity 52.8%; Pred. No. 6.9;
 Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
 QY 1 ATTATTTATTTTCATTAGATAGCCGGTTTTTTTACTACAACCTCAATAAGATGAACAGAAATG 60
 Db 116279 ATACATTTTATTCACAGACACTCAATTTTTTTTAAATATATATAAAATATGAATTAATG 116220
 QY 61 AATGGGTTACTGTCTTTTATAAGAGAGCTATAAAGATACATCATCATCTTCAGGCA 120
 Db 116219 AATTTAAAGTGAATATAATATAGAAATGTGCAACATTTCTATATATATATAAGTCAA 116160
 QY 121 TAAGGAGGAGAGATTTCAGCAACACGTGTGCTTACAAAGTGA 163
 Db 116159 CATACAGTGTCTAGCATATACAGCACACCCCTTACAAATGTA 116117
 RESULT 14
 BX537337
 LOCUS BX537337.9 GI:46406510
 DEFINITION Zebrafish DNA sequence from clone CH211-222E23 in linkage group 19,
 complete sequence.
 ACCESSION BX537337
 VERSION BX537337.9
 KEYWORDS HTG.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 217375)
 AUTHORS Woodmansey,R.
 TITLE Direct Submission
 JOURNAL Submitted (17-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish-help@sanger.ac.uk
 COMMENT On Apr 19, 2004 this sequence version replaced gi:46194230.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: zfish-help@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived
 zebrafish pUC subclones occasionally display inconsistency over the
 length of mononucleotide A/T runs and conserved TA repeats. Where
 this is found the longest good quality representation will be
 submitted.
 Repeat names beginning 'Dr' were identified by the Recon repeat
 discovery system (Zhikong Bao and Sean Eddy, submitted), and those
 beginning 'dr' were identified by Rick Waterman (Stephen Johnson
 lab, WashU). For further information see
 http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
 CH211-222E23 is from a CHORI-211 BAC library
 VECTOR: pTARBAC2.1.

```

FEATURES                                     Location/Qualifiers
source                                         1..217375
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-222E23"
/clone_lib="CHORI-211"

ORIGIN
Query Match                               11.8%; Score 39.8; DB 5; Length 217375;
Best Local Similarity 52.8%; Pred. No. 6.8;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 ATTATTATTTTCATTAGATACGCGGTTTTTACTACACTCAATGAATGACACAAATG 60
Db 157782 ATACATTTTATTCACACACTCAATTTTTTTTTTAAATATATATATAAATGATTAATG 157841

QY 61 AATGGGTAGTACTGTTTATAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA 120
Db 157842 AATTAAAGTGAATATATATATAGAAATGTGCAACATTTCTATATATTATTAAGTCAA 157901

QY 121 TAAGGGAGGGAGAGATTCAGCAACACAGTGTGCTTACAAGTGGGA 163
Db 157902 CATACAGTGCTCAGCATATATCAAGCACACCCCTTACAAGTGA 157944

RESULT 15
AC068698/c                                162605 bp      DNA      linear      HTG 24-AUG-2002
LOCUS                                     Homo sapiens chromosome 14, clone RP11-555P23 map 14, WORKING DRAFT
DEFINITION                               SEQUENCE, 13 unordered pieces.
AC068698
AC068698.2 GI:8569695
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE                                     Homo sapiens (human)
ORGANISM                                  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
Birren,B., Linton,L., Barna,N., Baetien,V., Bada,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meltrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 162605)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Baetien,V., Bada,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meltrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 162605)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Baetien,V., Bada,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,

```

```

Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meltrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 16, 2000 this sequence version replaced gi:771243.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9643
Center clone name: 555_P 23
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 154906 bases at least Q40
Consensus quality: 159118 bases at least Q30
Consensus quality: 160511 bases at least Q20
Insert size: 167000; agarose-Ep
Insert size: 161405; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-ep
Quality coverage: 4.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 464: contig of 464 bp in length
* 465 564: gap of 100 bp
* 565 2397: contig of 1833 bp in length
* 2398 2497: gap of 100 bp
* 2498 4504: contig of 2007 bp in length
* 4505 4604: gap of 100 bp
* 4605 10523: contig of 5919 bp in length
* 10524 10623: gap of 100 bp
* 10624 18758: contig of 8135 bp in length
* 18759 18858: gap of 100 bp
* 18859 27777: contig of 8919 bp in length
* 27778 27877: gap of 100 bp
* 27878 40462: contig of 12585 bp in length
* 40463 40562: gap of 100 bp
* 40563 53834: contig of 13272 bp in length
* 53835 53934: gap of 100 bp
* 53935 69869: contig of 15935 bp in length
* 69870 69970: gap of 100 bp
* 69971 87514: contig of 17545 bp in length
* 87515 103384: gap of 100 bp
* 103385 103484: contig of 15770 bp in length
* 103485 103484: gap of 100 bp

```

TITLE
JOURNAL

COMMENT

* 103485 129543: contig of 26059 bp in length
* 129544 129643: Gap of 100 bp
* 129644 162605: contig of 32962 bp in length.

FEATURES

source

1..162605
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/map="14"
/clone="RP11-555P23"
/clone_lib="RPC1-11 Human Male BAC"
1..464

misc_feature

/note="assembly_fragment"
clone_end:T7
vector_side:left"

misc_feature

565..7397
/note="assembly_fragment"

misc_feature

2498..4504
/note="assembly_fragment"

misc_feature

4605..10523
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

misc_feature

10624..18758
/note="assembly_fragment"

misc_feature

18859..27777
/note="assembly_fragment"

misc_feature

27878..40462
/note="assembly_fragment"

misc_feature

40563..53834
/note="assembly_fragment"

misc_feature

53935..69869
/note="assembly_fragment"

misc_feature

69970..87514
/note="assembly_fragment"

misc_feature

87615..103384
/note="assembly_fragment"

misc_feature

103485..129543
/note="assembly_fragment"

misc_feature

129644..162605
/note="assembly_fragment"

ORIGIN

Query Match 11.8%; Score 39.6; DB 2; Length 162605;
Best Local Similarity 51.7%; Pred. No. 7.9;
Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
Qy 18 ATAGCCGGTTTTTACTACAACTCAATAAGATGAACAGAAATGAATGGGTAGTACTGT 77
Db 104024 ACAGTCCCTTCCCCACTACAGGTCAACAAAAGGAGGAGGAAAAGCTTTGAGTTGGATA 103965
Qy 78 TTATAAGAGAGCTAATAAGATGACTATCATCTTTGAGGCAATAGGGAGGGAGAGATT 137
Db 103964 TGAGATTGAAGTTTGTGATAGCAAAATTACTATCTTTTAGTAAGTGAAGTCTTAGTTT 103905
Qy 138 CAGCAAAAGTGTGCTTACAAAGTGGAAAACAAAGTTAAACTAAAGTGACCCCT 191
Db 103904 CCATAAATAGGCTATTATTAAGTGAAGATGGAATAATGATGGGTCTTCCT 103851

Search completed: July 2, 2005, 17:52:56
Job time : 1532.96 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 10:20:51 ; Search time 196.01 Seconds

(without alignments)
10147.594 Million.cell updates/sec

Title: US-10-099-663-2

Perfect score: 336

Sequence: 1 attattatttttcattagata.....agacagaagaatggcattta 336

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	336	11 ADL90128	Adl90128 Chicken i
2	336	100.0	2381	11 ADL90127	Adl90127 Chicken i
3	39	11.6	2933	4 ABL18616	Ab118616 Drosophil
4	37.2	11.1	2927	4 ABL05564	Ab105564 Drosophil
5	37.2	11.1	123526	10 ADU79962	Adj79962 Human gli
6	36.6	10.9	403	4 AAS58239	Aas58239 cDNA #915
7	36.6	10.9	3322	4 ABL10846	Ab110846 Drosophil
8	36.6	10.9	53585	2 AAX20251	Aax20251 Borrelia
9	36.2	10.8	1587	4 AAH53268	Aah53268 S. epider
10	36.2	10.8	1878	6 ABN92204	Abn92204 Staphyloc
11	36.2	10.8	1878	13 ADS01975	Ads01975 Staphyloc
12	36.2	10.8	3000	4 AAH54543	Aah54543 S. epider
13	36.2	10.8	3368	4 AAH54805	Aah54805 S. epider
14	36.2	10.8	6088	2 AAX84331	Aax84331 Stealth v
15	36.2	10.8	8648	6 ABK31354	Abk31354 Signal tr
16	36.2	10.8	8648	6 ABL70573	Ab170573 Chemicall
17	35.8	10.7	714	6 ABL56240	Ab156240 AmEPV pho
18	35.8	10.7	50000	6 ABL56202	Ab156202 AmEPV gen
19	35.2	10.5	663	2 ADR01495	Adr01495 A. gossyp
20	35.2	10.5	696	2 ADR02087	Adr02087 A. gossyp

c 21	35.2	10.5	719	2	ADR02446	Adr02446 A. gossyp
c 22	35.2	10.5	856	2	ADR01597	Adr01597 A. gossyp
c 23	35.2	10.5	8622	6	ABL34142	Ab134142 Human imm
c 24	35.2	10.5	15714	6	ABL33173	Ab133173 Human imm
c 25	35.2	10.5	15714	6	ABQ67058	Abq67058 Human ang
c 26	35	10.4	2839	4	AAH62755	Aah62755 Shrimp wh
c 27	35	10.4	134738	11	ACN44182	Acn44182 Human gen
c 28	35	10.4	305107	4	AAHG2689	Aah62689 Shrimp wh
c 29	34.8	10.4	18488	6	ABA01444	AbA01444 Streptoco
c 30	34.8	10.4	19738	6	ABA01436	AbA01436 Streptoco
c 31	34.8	10.4	28564	10	ADD47140	AdD47140 Human gen
c 32	34.8	10.4	75899	6	ABK85261	Abk85261 Human gen
c 33	34.8	10.4	75899	12	ADL13990	Adl13990 Human pro
c 34	34.6	10.3	6072	6	ABL32031	Ab132031 Human imm
c 35	34.4	10.2	1137	10	ADC93219	Adc93219 E. faeciu
c 36	34.4	10.2	1173	13	ADS46350	AdS46350 Bacterial
c 37	34.4	10.2	10957	6	ABL33111	Ab133111 Human imm
c 38	34.4	10.2	12751	4	AAK85146	Aak85146 Human imm
c 39	34.4	10.2	12751	4	AAK85147	Aak85147 Human imm
c 40	34.4	10.2	69727	10	ACF65374	Acf65374 Phototrab
c 41	34.4	10.2	110000	2	AAV21209_15	Continuation (16 o
c 42	34.4	10.2	110000	10	ACF67367_35	Continuation (36 o
c 43	34.2	10.2	400	4	AAI80079	Aai80079 Human pol
c 44	34.2	10.2	7401	12	ADO26270	Ado26270 C35 promo
c 45	34.2	10.2	8605	6	ABL32535	Ab132535 Human imm

ALIGNMENTS

RESULT 1

ADL90128
ID ADL90128 standard; DNA; 336 BP.

XX ADL90128;

XX 20-MAY-2004 (first entry)

XX Chicken intestinal fatty acid binding protein, iFABP, gene, promoter.

XX Chicken; ds; intestinal fatty acid binding protein; iFABP;

XX gut specific promoter; transgenic; promoter.

XX Gallus gallus.

XX US2003177516-A1.

XX 18-SEP-2003.

XX 14-MAR-2002; 2002US-00099663.

XX 14-MAR-2002; 2002US-00099663.

XX (HORS/) HORSEMAN N D.

XX (PRAT/) PRATT S L.

XX Horseman ND, Pratt SL;

XX WPI; 2003-898653/82.

XX New nucleic acid molecule comprising an isolated avian gut-specific gene expression control region, useful for regulating heterologous nucleic acids in transgenic avians, and for generating transgenic birds.

XX Claim 1; SEQ ID NO 2; 28pp; English.

XX The invention relates to an isolated nucleic acid comprising an isolated avian gut-specific gene expression control region appearing as ADL90127 (Chicken intestinal fatty acid binding protein, iFABP, gene, 5' region or ADL90128 (Chicken iFABP promoter) or its degenerate variant. Also included are a recombinant DNA molecule comprising an isolate avian gut-specific gene expression control region operably linked to a nucleic acid insert encoding a polypeptide, an expression vector that integrates

CC into a host cell (and comprising the isolated avian gut-specific gene
 CC expression control region), expressing a heterologous polypeptide in a
 CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
 CC molecule, and culturing the transfected cell in a medium suitable for
 CC expression of a heterologous polypeptide under the control of an avian
 CC intestinal fatty acid binding protein (iFABP) or cp35 gene expression
 CC control region encoded by the recombinant DNA molecule), a eukaryotic
 CC cell transformed with the expression vector (or its progeny, which
 CC expresses a heterologous polypeptide) and a transgenic avian having a
 CC heterologous polynucleotide sequence comprising the nucleic acid insert.
 CC The nucleic acids are useful for regulating heterologous nucleic acids in
 CC transgenic avians, as probes in nucleic acid hybridisation assays for
 CC detecting the iFABP gene expression control region, and for generating
 CC transgenic birds. The present sequence is the Chicken intestinal fatty
 CC acid binding protein, iFABP, gene, promoter.

XX SQ Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 11; Length 336;
 Best Local Similarity 100.0%; Pred. No. 2e-87;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTATTATTTTCATTAGATAGCGGTTTTTACTACAACCTCAAATAAGATGAACAGATG 60
 DB 1 ATTATTATTTTCATTAGATAGCGGTTTTTACTACAACCTCAAATAAGATGAACAGATG 60
 QY 61 AATGGGTTAGTACTGTTTATAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA 120
 DB 61 AATGGGTTAGTACTGTTTATAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA 120
 QY 121 TAAGGAGGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAAAACAAGTTAAACTAAA 180
 DB 121 TAAGGAGGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAAAACAAGTTAAACTAAA 180
 QY 181 GTGACCCCTCTCTTGACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
 DB 181 GTGACCCCTCTCTTGACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
 QY 241 TGTAAATTTGCTTCTTGATAGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTA 300
 DB 241 TGTAAATTTGCTTCTTGATAGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTA 300
 QY 301 CCAGAAAGTCTGCCCTACAGACAGAAAGATGGCATTTA 336
 DB 301 CCAGAAAGTCTGCCCTACAGACAGAAAGATGGCATTTA 336

RESULT 2
 ADL90127
 ID ADL90127 standard; DNA; 2381 BP.

XX AC ADL90127;
 XX XX
 XX 20-MAY-2004 (first entry)

XX DE Chicken intestinal fatty acid binding protein, iFABP, gene, 5' region.
 XX XX
 XX KW Chicken; ds; intestinal fatty acid binding protein; iFABP;
 XX KW gut specific promoter; transgenic.

XX OS Gallus gallus.
 XX XX
 XX PN US2003177516-A1.

XX PD 18-SEP-2003.
 XX XX
 XX PF 14-MAR-2002; 2002US-00099663.
 XX XX
 XX PR 14-MAR-2002; 2002US-00099663.

XX XX (HORS/) HORSEMAN N D.
 XX PA (PRAT/) PRATT S L.
 XX XX

PI Horseman ND, Pratt SL;
 XX DR WPI; 2003-898553/82.
 XX PT New nucleic acid molecule comprising an isolated avian gut-specific gene
 XX PT expression control region, useful for regulating heterologous nucleic
 XX PT acids in transgenic avians, and for generating transgenic birds.
 XX PS Claim 1; SEQ ID NO 1; 28pp; English.

XX CC The invention relates to an isolated nucleic acid comprising an isolated
 XX CC avian gut-specific gene expression control region appearing as
 XX CC ADL90127(Chicken intestinal fatty acid binding protein, iFABP, gene, 5'
 XX CC region or ADL90128 (chicken iFABP promoter) or its degenerate variant.
 XX CC Also included are a recombinant DNA molecule comprising an isolate avian
 XX CC gut-specific gene expression control region operably linked to a nucleic
 XX CC acid insert encoding a polypeptide, an expression vector that integrates
 XX CC into a host cell (and comprising the isolated avian gut-specific gene
 XX CC expression control region), expressing a heterologous polypeptide in a
 XX CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
 XX CC molecule, and culturing the transfected cell in a medium suitable for
 XX CC expression of a heterologous polypeptide under the control of an avian
 XX CC intestinal fatty acid binding protein (iFABP) or cp35 gene expression
 XX CC control region encoded by the recombinant DNA molecule), a eukaryotic
 XX CC cell transformed with the expression vector (or its progeny, which
 XX CC expresses a heterologous polypeptide) and a transgenic avian having a
 XX CC heterologous polynucleotide sequence comprising the nucleic acid insert.
 XX CC The nucleic acids are useful for regulating heterologous nucleic acids in
 XX CC transgenic avians, as probes in nucleic acid hybridisation assays for
 XX CC detecting the iFABP gene expression control region, and for generating
 XX CC transgenic birds. The present sequence is the Chicken intestinal fatty
 XX CC acid binding protein, iFABP, gene, 5' region.

SQ Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 336; DB 11; Length 2381;
 Best Local Similarity 100.0%; Pred. No. 3.8e-87;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTATTATTTTCATTAGATAGCGGTTTTTACTACAACCTCAAATAAGATGAACAGATG 60
 DB 1301 ATTATTATTTTCATTAGATAGCGGTTTTTACTACAACCTCAAATAAGATGAACAGATG 1360
 QY 61 AATGGGTTAGTACTGTTTATAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA 120
 DB 1361 AATGGGTTAGTACTGTTTATAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA 1420
 QY 121 TAAGGAGGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAAAACAAGTTAAACTAAA 180
 DB 1421 TAAGGAGGGAGAGATTTCAGCAACAGTGTGCTTACAGTGGAAAAACAAGTTAAACTAAA 1480
 QY 181 GTGACCCCTCTCTTGACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
 DB 1481 GTGACCCCTCTCTTGACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 1540
 QY 241 TGTAAATTTGCTTCTTGATAGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTA 300
 DB 1541 TGTAAATTTGCTTCTTGATAGCTGTTTCATAAATTTCTTTGCAAGCTCTGCTACTTA 1600
 QY 301 CCAGAAAGTCTGCCCTACAGACAGAAAGATGGCATTTA 336
 DB 1601 CCAGAAAGTCTGCCCTACAGACAGAAAGATGGCATTTA 1636

RESULT 3
 ABL18616
 ID ABL18616 standard; DNA; 2933 BP.

XX AC ABL18616;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 7321.

```
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 7321; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2933 BP; 883 A; 628 C; 606 G; 816 T; 0 U; 0 Other;
Query Match 11.6%; Score 39; DB 4; Length 2933;
Best Local Similarity 52.1%; Pred. No. 0.75;
Matches 87; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 73 ACTGTTTATAAGAGAGTAAAGATACATCATCTATCATCTATTGAGGCAATAGGGAGGAG 132
DB 2766 ACATTAATAATGCAGATCTTACTTAAACGAATAATAGGATTATGTCGAGAAGAGGAGAT 2825
QY 133 AGATTACAGCAAAACAGTGTGCTTACAAAGTGAAACAAAGTTAAACTAAAGTGACCCCTC 192
DB 2826 ATATTACAGAAAATTTCTTATCGATGTGAACAAAGTGAACCGAAATTTGCATTCATC 2885
QY 193 CTTGACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCCATCATC 239
DB 2886 CGTGGCCCACTGAATGCAATCACTTAGTCGACACCAACCACC 2932
RESULT 4
ABLO6564
ID ABL06564 standard; cDNA; 2927 BP.
XX XX
XX AC ABL06564;
XX XX
XX DT 26-MAR-2002 (first entry)
XX XX
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 14174.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
```

```
XX XX
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR P-PSDB; ABB62461.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 14174; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2927 BP; 882 A; 623 C; 606 G; 816 T; 0 U; 0 Other;
Query Match 11.1%; Score 37.2; DB 4; Length 2927;
Best Local Similarity 51.9%; Pred. No. 2.5;
Matches 84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 73 ACTGTTTATAAGAGAGTAAAGATACATCATCTATCATCTATTGAGGCAATAGGGAGGAG 132
DB 2766 ACATTAATAATGCAGATCTTACTTAAACGAATAATAGGATTATGTCGAGAAGAGGAGAT 2825
QY 133 AGATTACAGCAAAACAGTGTGCTTACAAAGTGAAACAAAGTTAAACTAAAGTGACCCCTC 192
DB 2826 ATATTACAGAAAATTTCTTATCGATGTGAACAAAGTGAACCGAAATTTGCATTCATC 2885
QY 193 CTTGACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCCAC 234
DB 2886 CGTGGCCCACTGAATGCAATCACTTAGTCGACACCAACCACATA 2927
RESULT 5
ADJ79962/c
ID ADJ79962 standard; DNA; 123526 BP.
XX XX
XX AC ADJ79962;
XX XX
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE Human glioma-associated oncogene-3 related DNA, SEQ ID No 11.
XX KW glioma-associated oncogene-3; GAO3; cytostatic; developmental disorder;
XX KW Greig's cephalopolysyndactyly; Pallister-Hall syndrome;
XX KW post-axial polydactyly; holoprosencephaly; Rubenstein-Teybi syndrome;
XX KW basal cell nevoid syndrome; hyperproliferative disorder; cancer; human;
XX KW ds.
XX OS Homo sapiens.
XX XX
XX PN WO2003008549-A2.
XX PD 30-JAN-2003.
XX XX
```

```
PF 15-JUL-2002; 2002WO-US022630.
XX
XX
PR 18-JUL-2001; 2001US-00910185.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Bennett FC, Freier SM;
XX
XX WPI; 2003-239322/23.
XX
XX New antisense oligonucleotides targeted to a nucleic acid encoding glioma
PT -associated oncogene-3, useful for treating developmental disorders (e.g.
PT holoprosencephaly) and hyperproliferative disorders (e.g. cancer).
XX
XX Disclosure; SEQ ID NO 11; 175pp; English.
XX
XX The invention relates to a novel compound 8-50 nucleobases in length
CC targeted to a nucleic acid encoding glioma-associated oncogene-3 (GAO3)
CC or a splice variant of GAO3. The novel compound specifically hybridizes
CC with and inhibits the expression of GAO3 or its splice variant, or
CC specifically hybridizes with an 8-nucleobase portion of an active site on
CC a nucleic acid encoding GAO3. The antisense compound has cytostatic
CC activity. The antisense compound is useful for treating a disease or
CC condition associated with glioma-associated oncogene-3 (GAO3), such as a
CC developmental disorder including Greig's cephalopolysyndactyly, Pallister
CC -Hall syndrome, post-axial polydactyly, holoprosencephaly, Rubenstein-
CC teybi syndrome or basal cell nevoid syndrome, and a hyperproliferative
CC disorder, such as cancer. This polynucleotide represents a DNA sequence
CC relating to the human glioma-associated oncogene-3 (GAO3) of the
CC invention.
XX
XX Sequence 123526 BP; 34368 A; 24882 C; 26026 G; 38250 T; 0 U; 0 Other;
SQ
Query Match 11.1%; Score 37.2; DB 10; Length 123526;
Best Local Similarity 53.4%; Pred. No. 9;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 12 CATTAGATACCCGGTTTTTTTACTACAACTCAATAAGATGAACAGAAATGAATGGGTAGT 71
Db 107264 CATTTCCTCCCAATGCCCTTTTAATTAAATGCTACAGAAATGCACAGAGAGAGGGGTGCT 107205
QY 72 GACTGTTTATAAGAGAGAGTAATAAGATACTATCATCTTTGAGGCAATAAGGGAGGGA 131
Db 107204 GTGTACCCATTAAATAATGAGATGAGAGAAAGAGAGAAAGGCGAGAGGGAGGA 107145
QY 132 GAGATTGAGCAAAACAGTGTGCTTACA 157
Db 107144 GGGAGAAACCCGAAAGCATCCATGCA 107119
RESULT 6
AAS58239
ID AAS58239 standard; cDNA; 403 BP.
XX
XX AAS58239;
XX
XX 13-FEB-2002 (first entry)
XX
XX cDNA #915 encoding portion of a human colon tumour protein.
XX
XX Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
XX Homo sapiens.
XX
XX WO200173027-A2.
XX
XX 04-OCT-2001.
XX
XX 22-MAR-2001; 2001WO-US009246.
XX
XX 24-MAR-2000; 2000US-0191597P.
XX
XX 04-MAY-2000; 2000US-0202024P.
XX
XX 05-MAY-2000; 2000US-0202189P.
```

```
XX
XX (CORI-) CORIXA CORP.
XX
XX Meagher MJ, Xu J, King GE;
XX
XX WPI; 2001-611627/70.
XX
XX New colon tumor proteins and related nucleic acid, useful for treatment,
PT prevention, diagnosis and monitoring of cancer.
XX
XX Claim 4; Page 202; 299pp; English.
XX
XX Th present invention relates to the isolation of novel cDNA sequences
CC encoding for at least an immunogenic portion of human colon tumour
CC proteins. The sequences of the invention are useful in pharmaceutical
CC compositions and vaccines for the prevention and treatment of cancers
CC such as colon cancer. They are also useful for the diagnosis and
CC monitoring of such cancers. Antibodies to the colon tumour proteins and
CC antigen presenting cells that express polynucleotides encoding colon
CC tumour proteins can be used to inhibit the development of cancers. The
CC cells that react specifically with colon tumour proteins are useful for
CC removing tumour cells from samples (e.g. blood) and for cancer treatment.
CC The polynucleotides sequences are also useful in gene therapy. AAS57325-
CC AAS58880 represent the cDNA sequences of the invention that encode for
CC portions of human colon tumour proteins
XX
XX Sequence 403 BP; 136 A; 63 C; 88 G; 105 T; 0 U; 11 Other;
SQ
Query Match 10.9%; Score 36.6; DB 4; Length 403;
Best Local Similarity 50.6%; Pred. No. 1.9;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 26 TTTTCTTACTACAACTCAATAAGATGAACAGAAATGAATGGGTAGTACTGTTTATAAAG 85
Db 178 TGTGTTAAAGTTTAAAGAAAAAGAGCTGCAGAGATATTTATAAACTGCTTTTAGAAAAA 237
QY 85 AAGAGTATTAAGATACTATCATCTTTGAGGCAATAAGGGAGGAGAGATTTCAGCAAC 145
Db 238 AACAAGCAGAGAGACCATTTGACCATATGATGGAAAGGGGAGAAAGTATTATAGAAAC 297
QY 146 AGTGTGCTTACAAAGTGGAAAAACAAGTTAAACTTAAAGTGACCCCCCT 191
Db 298 TTTGCTAGTTNAAAAAAGAAAAAAGAAAAAAGAACTTNGCNGAACCCCT 343
RESULT 7
ABL10846
ID ABL10846 standard; cDNA; 3322 BP.
XX
XX ABL10846;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 27020.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
```

XX	Claim 1; Page 801-831; 1128pp; English.
XX	
XX	AAH520248 to AAH520402 represent polynucleotide sequences isolated from
CC	<i>Borrelia burgdorferi</i> (Bb). Products derived from Bb can be used for the
CC	detection, diagnosis, characterisation, prevention and therapy of Bb
CC	infections, e.g. Lyme disease. They can also be used for the production
CC	of biosynthetic products, e.g. enzymes. <i>Borrelia</i> belongs to a family of
CC	motile, spiral-shaped bacteria called Spirochetes. Spirochetes are
CC	pathogenic in humans and <i>Borrelia</i> causes epidemic and endemic relapsing
CC	fever, and Lyme borreliosis, more commonly known as Lyme disease
XX	
XX	Sequence 53585 BP; 19235 A; 8081 C; 7001 G; 19266 T; 0 U; 2 Other;
XX	
XX	Query Match 10.9%; Score 36.6; DB 2; Length 53585;
XX	Best Local Similarity 58.9%; Pred. No. 10;
XX	Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
XX	
Qy	78 TTATAAGAAGAGTATTAAGAATATCTATCATCTTTGAGGCAATAAGGGAGGAGAGATT 137
Db	44978 TTAATAATGTTAATAATAAGATTACTCACTTATATCAAGCATTAAATGATGAGAGAGATG 44919
XX	
Qy	138 CAGCAACAGTGTGCTTACAGTGGAAACAGTTAAACTAACTGA 184
Db	44918 CAAGCATCATTTGAACATTTAAGAGAAGAGAGAAATGAACTAATGTCA 44872
XX	
XX	RESULT 9
XX	AAH53268
ID	AAH53268 standard; DNA; 1587 BP.
AC	AAH53268;
XX	
DT	03-SEP-2001 (first entry)
XX	
DE	S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1929.
XX	
KW	<i>Staphylococcus epidermidis</i> SR1 strain; infection; diagnosis; vaccination;
KW	endocarditis; ds.
XX	
OS	<i>Staphylococcus epidermidis</i> .
XX	
PN	WO200134809-A2.
XX	
PD	17-MAY-2001.
XX	
PF	09-NOV-2000; 2000WO-US030782.
XX	
PR	09-NOV-1999; 99US-0164258P.
XX	
PA	(GLAX) GLAXO GROUP LTD.
XX	
PI	Kimmerly WJ;
XX	
WPI	WPI; 2001-316495/33.
DR	P-PSDB; AAH82418.
XX	
PT	Nucleic acids encoding polypeptides from <i>Staphylococcus epidermidis</i> ,
XX	useful for vaccinating against infections, e.g. endocarditis.
XX	
XX	Claim 8; Page 527-528; 2188pp; English.
XX	
XX	AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC	(II), given in AAH81454 to AAH83120, from <i>Staphylococcus epidermidis</i> . (I
CC	and (II) can have antibacterial activity and therefore can be used in
CC	vaccination. The nucleic acids (I) may be used to produce the S.
CC	epidermidis polypeptides (II) via the production of vectors containing
CC	them which are used to produce hosts cells which express the
CC	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC	used to vaccinate subjects and to raise antibodies against the bacteria.
CC	The polypeptides may also be used to assay for other inhibitors of their
CC	activity and therefore identify compounds that may be used for the
CC	treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to

them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH53990 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention. N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464

[illegible]

RESULT 13
 AAH54805/c
 ID AAH54805 standard; DNA; 3368 BP.
 XX
 AC AAH54805;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:4169.
 XX
 KW Staphylococcus epidermidis SFL strain; infection; diagnosis; vaccination;
 KW endocarditis; ds.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US030782.
 XX
 PR 09-NOV-1999; 99US-0164358P.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Kimmerly WJ;
 XX
 DR WPI; 2001-316495/33.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis.
 PS Claim 8; Page 1875-1876; 2188pp; English.

XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 13:46:26 ; Search time 61.0909 Seconds
(without alignments)
8999.522 Million cell updates/sec

Title: US-10-099-663-2
Perfect score: 336
Sequence: 1 attattatttccattagata.....agacagaagatggcattta 336

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.2	11.1	1141	US-09-806-708B-22	Sequence 22, Appl
c	36.8	11.0	43117	US-09-949-016-17589	Sequence 17589, A
3	36.2	10.8	1587	US-09-710-279-1929	Sequence 1929, Ap
4	36.2	10.8	1878	US-09-134-001C-1667	Sequence 1667, Ap
5	36.2	10.8	3000	US-09-710-279-3907	Sequence 3907, Ap
c	36.2	10.8	3368	US-09-710-279-4169	Sequence 4169, Ap
7	36	10.7	168394	US-09-949-016-13002	Sequence 13002, A
8	35.8	10.7	7218	US-08-232-463-14	Sequence 14, Appl
c	35.6	10.6	601	US-09-949-016-44395	Sequence 44395, A
10	35.6	10.6	601	US-09-949-016-58898	Sequence 58898, A
11	35.6	10.6	139562	US-09-949-016-13451	Sequence 13451, A
c	35.6	10.6	256171	US-09-949-016-12822	Sequence 12822, A
13	35.6	10.6	256176	US-09-949-016-15224	Sequence 15224, A
14	35.2	10.5	601	US-09-949-016-134974	Sequence 134974, A
c	35.2	10.5	663	US-08-998-416-187	Sequence 187, App
16	35.2	10.5	696	US-08-998-416-1779	Sequence 779, App
c	35.2	10.5	719	US-08-998-416-1138	Sequence 1138, App
18	35.2	10.5	856	US-08-998-416-289	Sequence 289, App
19	35	10.4	399	US-09-621-976-8976	Sequence 8976, Ap
c	34.8	10.4	78125	US-09-949-016-16006	Sequence 16006, A
21	34.6	10.3	601	US-09-949-016-173718	Sequence 173718, A
22	34.6	10.3	601	US-09-949-016-173765	Sequence 173765, A
23	34.6	10.3	81819	US-09-949-016-16661	Sequence 16661, A
24	34.6	10.3	81819	US-09-949-016-16662	Sequence 16662, A
25	34.4	10.2	1137	US-09-107-532A-2846	Sequence 2846, Ap
c	34.4	10.2	1664976	US-08-916-421B-1	Sequence 1, Appli
27	34.4	10.2	1664976	US-09-692-570-1	Sequence 1, Appli

28	34.2	10.2	98844	3	US-09-791-211-10	Sequence 10, Appl
29	34.2	10.2	143776	4	US-09-949-001-29	Sequence 29, Appl
30	34.2	10.2	144034	4	US-09-949-001-35	Sequence 35, Appl
c	33.6	10.0	2119	3	US-09-240-639-7	Sequence 7, Appli
31	33.6	10.0	2119	4	US-09-908-510A-7	Sequence 7, Appli
c	33.6	10.0	2119	4	US-09-905-744B-7	Sequence 7, Appli
32	33.6	10.0	2119	4	US-10-107-576-7	Sequence 7, Appli
c	33.6	10.0	2119	4	US-09-905-732B-7	Sequence 7, Appli
33	33.6	10.0	2119	4	US-09-905-743B-7	Sequence 7, Appli
c	33.6	10.0	2119	4	US-09-949-016-13310	Sequence 13310, A
34	33.4	9.9	31385	4	US-09-949-016-12358	Sequence 12358, A
c	33.4	9.9	73788	4	US-08-178-242-4	Sequence 4, Appli
40	33	9.8	2642	1	US-08-955-091-4	Sequence 4, Appli
41	33	9.8	2642	2	US-09-225-510-4	Sequence 4, Appli
42	33	9.8	2642	3	US-08-178-242-14	Sequence 14, Appl
43	33	9.8	2870	1	US-08-955-091-14	Sequence 14, Appl
44	33	9.8	2870	2	US-09-225-510-14	Sequence 14, Appl
45	33	9.8	2870	3		

ALIGNMENTS

RESULT 1

US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAEl promoters
US-09-806-708B-22

Query Match	11.1%;	Score	37.2;	DB	4;	Length	1141;
Best Local Similarity	10.7%;	Pred. No.	0.24;				
Matches	31;	Conservative	123;	Mismatches	135;	Indels	0;
Gaps	0;						
QY	1	ATTATATTTCATTAGATAGCGGTTTCTTACTCACTCAATCAATGAATGAACAATG	60				
Db	285	AKAMCKRAKYGNWRABVNSTCTTWSKTKTKVTSWANNCRAGDANKDKHKKWWSAAMG	344				
QY	61	AATGGCTAGTCACTGTTTATAAAGAGAGTAATAAGATACATCATCATTTTGAGGCAA	120				
Db	345	VYNNNNNNWTKYKARHARDWVHSAWKHANAHAHYSKKTWBYKRTWVNNNGT	404				
QY	121	TAAGGGAGGAGAGATTACGCAACAGTGTGTTCAAGTGGAACAAACAAGTTAACTAA	180				
Db	405	TWKKRWAWYKMDWDBGTNNNNNGRTYYGWTNKKQWTVYKWKANNCKWRWDHKT	464				
QY	181	GTGACCCCTCTTGCAAGATCAATGCCAGTTGACCTTTAGCCACCATCATCA	240				
Db	465	CTHNNNTWMMKTYNNNNYNNKSHRBAAYVTWYNNRRYAHANNNDWYWKA	524				
QY	241	TGTAAATGCTTTCCTGATAAGCCTGTTTCATAAATCTCTTTGCAAGC	289				
Db	525	CTWYKYBVCCKWNNYAAWYTKSSWNYTSRYRWKNNNSWRSDTRSM	573				

RESULT 2

US-09-949-016-17589/c
; Sequence 17589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17589
; LENGTH: 43117
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17589

Query Match 11.0%; Score 36.8; DB 4; Length 43117;
Best Local Similarity 52.6%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 30 TTACTCAAACTCAATAGATGAACAGAAATGAATGGTTAGTGACTGTTTATAAAGAGA 89
Db 20807 TGAACAACAATATATGTTGATGAAAAATTATTGAGGATTAATAAAGATGATAAAATGCA 20748
QY 90 GTAATGAAGATCTATCATCTTTGAGCAATTAAGGAGGAGAGAGATTGAGCAAAACAGTG 149
Db 20747 CCTTAAGAGACAAAATAGTAAGTTAGGAAATAGCCTCTGGCGGAATCTGAAGACAGGA 20688
QY 150 TGCCTCAAGTGGAAAACAAGTTAAACTAAAG 181
Db 20687 GTTTTAAAGATGACCAACCAATTAGAATGAAG 20656

RESULT 3
US-09-710-279-1929
; Sequence 1929, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1929
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1929

Query Match 10.8%; Score 36.2; DB 4; Length 1587;
Best Local Similarity 53.1%; Pred. No. 0.54; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 141 CAAACAGTGTGCTTACAGTGGAAAACAAGTTAACTAAAGTGACCCCTCCTTGACAA 200
Db 203 CAGACGATGACTTTAAAGTTGAAAACCTTTATTTTACAGAGTTACCTCAAGCCCATCTA 262
QY 201 GATCAATGCCAGTTGAGCTTTAGCCACCATCATCATGTAATTCCTTCTTGATA 260

Db 263 GAAAAATAAAACAAATATATCTCAAGCATATAGATAATCATCAATATCAACTTCTGAAA 322
QY 261 AGCCTGTTCAATAAATCTCTTTGCA 285
Db 323 ATAAAAATACAACAATTAATTGTGCA 347
RESULT 4
US-09-134-001C-1667
; Sequence 1667, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1667
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1667

Query Match 10.8%; Score 36.2; DB 3; Length 1878;
Best Local Similarity 53.1%; Pred. No. 0.58; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 141 CAAACAGTGTGCTTACAAAGTGGAAAACAAGTTAAACTAAAGTGACCCCTCCTTGACAA 200
Db 494 CAGACGATGACTTTAAAGTTGAAAACCTTATTTTACAGAGTTACCTCAAGCCCATCTA 553
QY 201 GATCAATGCCAGTTGAGCTTTAGCCAGCCACATCATCATGTAATTCCTTCTTGATA 260
Db 554 GAAAAATAAAACAAATATATCTCAAGCATATAGATAATCATCAATATCAACTTCTGAAA 613
QY 261 AGCCTGTTCAATAAATCTCTTTGCA 285
Db 614 ATAAAAATACAACAATTAATTGTGCA 638

RESULT 5
US-09-710-279-3907
; Sequence 3907, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3907
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3907

Query Match 10.8%; Score 36.2; DB 4; Length 3000;
Best Local Similarity 53.1%; Pred. No. 0.69; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Query Match
10.7%; Score 35.8; DB 1; Length 7218;

```
Best Local Similarity 7.4%; Pred. No. 1.3;
Matches 13; Conservative 100; Mismatches 62; Indels 0; Gaps 0;

QY 10 TTCAATGATAGCGGGTTTCTTACTCAACTCAATAAGATGAACAGATGAATGGGTTA 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1459 TTAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400

QY 70 GTGACTGTTTATAAAGAGATGAATAAGATCATCTATCATCTTTGAGGCAATAAGGAGG 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1399 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1340

QY 130 GAGAGATTTCAGCAACAGTGTGCTTACAAGTGGAACAAAGTTAAACTAAAGTGA 184
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1339 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1285

RESULT 9
US-09-949-016-44395/c
; Sequence 44395, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44395
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-44395

Query Match 10.6%; Score 35.6; DB 4; Length 601;
Best Local Similarity 57.4%; Pred. No. 0.57;
Matches 62; Conservative 1; Mismatches 45; Indels 0; Gaps 0;

QY 27 TTTTACTCAACTCAATAAGATGAACAGATGAATGGTGTAGTGACTGTTTATAAGA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 TTTTGACCTCAATAAGATGAATCAATAGATAAAGAAAATAGTTGCTCTTTATAGCT 339

QY 87 AGAGTAATAAGATCATCTATCATCTTTGAGGCAATAAGGAGGAGAG 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 TTATTTATATTATTAATTTCTCAACAATGAGTAGTAATAAGGAGTCAGAG 291

RESULT 10
US-09-949-016-58898
; Sequence 58898, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 58898
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-58898

Query Match 10.6%; Score 35.6; DB 4; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.57;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 43 AATAAGATCAACAGATGAATGGTGTAGTGACTGTTTATAAAGAGAGTAATAAGATAC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 ATTACATGAGATGCACTGATGACCTAATGCTTTTGAATATTATACCAATGAAAGAT 435

QY 103 TATCATCAFTTTCAGGCAATAAGGGAGAGATTCAGCAAAACAGTGTCTTTACAAGTGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 TAATATGATGCCAGGAAGTAACTGACCAAGAGCAGAGAAAGAGAGGAGGCACAAAGTTC 495

QY 163 AAAACAAGTTAACTAAAGTGACCCC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 AAAAGTCTGTAGAGGCCAGAGAGACC 521

RESULT 11
US-09-949-016-13451
; Sequence 13451, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13451
; LENGTH: 139562
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(139562)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-949-016-13451

Query Match 10.6%; Score 35.6; DB 4; Length 139562;
Best Local Similarity 52.7%; Pred. No. 4.6;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 43 AATAAGATCAACAGATGAATGGTGTAGTGACTGTTTATAAAGAGAGTAATAAGATAC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38132 ATTTACATGGACTGCAGTGATGACCTAATGCTTTTGAATATTATACCAATGAAAGAT 38191

QY 103 TATCATCAFTTTCAGGCAATAAGGGAGGAGATTCAGCAAAACAGTGTCTTACAAGTGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38192 TAATATGATGCCAGGAAGTAACTGACCAAGAGCAGAGAAAGAGAGGAGGCACAAAGTTC 38251

QY 163 AAAACAAGTTAACTAAAGTGACCCC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38252 AAAAGTCTGTAGAGGCCAGAGAGACC 38277

RESULT 12
US-09-949-016-12822/c
; Sequence 12822, Application US/09949016
; Patent No. 6812339
```

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12822
; LENGTH: 256171
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(256171)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12822

Query Match      10.6%; Score 35.6; DB 4; Length 256171;
Best Local Similarity 49.0%; Pred. No. 5.7;
Matches 95; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 4 ATTATTTTCATTAGATAGCGGTTTTTACTACAACTCAAAATAGATGAACAGAAATGAAT 63
DB 88458 ACTGATTTTCATCATTTACACACCTTACATACATAAGAAAATATCACATGAATCTCATATAAT 88399

QY 64 GGGTTAGTACTGTTTATATAAGAGAGTATAAAGATACATCATCATCTTTTGAGGCAATAA 123
DB 88398 ATGTAAATATTTGTGTATCAATTTTAAAAAACCCAAAAAATTTTCTTATGATCATGTAA 88339

QY 124 GCGAGGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTTAAAGTG 183
DB 88338 ATAATACATACCCATCGTTTAAAGTTTGGAAATTCACAGAAGACATGATGAAAAAATTA 88279

QY 184 ACCCCCTCCTTGA 197
DB 88278 AAATACCCATTAA 88265

RESULT 14
US-09-949-016-134974
; Sequence 134974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134974
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-134974

Query Match      10.5%; Score 35.2; DB 4; Length 601;
Best Local Similarity 48.5%; Pred. No. 0.75;
Matches 94; Conservative 1; Mismatches 99; Indels 0; Gaps 0;

QY 4 ATTATTTTCATTAGATAGCGGTTTTTACTACAACTCAAAATAGATGAACAGAAATGAAT 63
DB 232 ACTGATTTTCATCATTTACACACCTTACATACATAAGAAAATATCACATGAATCTCATATAAT 291

QY 64 GGGTTAGTACTGTTTATATAAGAGAGTATAAAGATACATCATCATCTTTTGAGGCAATAA 123
DB 292 ATGTAAATATTTGTGTATCAATTTTAAAAAACCCAAAAAATTTTCTTATGATCATGTAA 351

QY 124 GCGAGGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTTAAAGTG 183
DB 352 ATAATACATACCCATCGTTTAAAGTTTGGAAATTCACAGAAGACATGATGAAAAAATTA 411

QY 184 ACCCCCTCCTTGA 197
DB 412 AAATACCCATTAA 425

RESULT 15
US-08-998-416-187/c
; Sequence 187, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15524
; LENGTH: 256176
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(256176)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15524/c

Query Match      10.6%; Score 35.6; DB 4; Length 256171;
Best Local Similarity 49.0%; Pred. No. 5.7;
Matches 95; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 4 ATTATTTTCATTAGATAGCGGTTTTTACTACAACTCAAAATAGATGAACAGAAATGAAT 63
DB 88458 ACTGATTTTCATCATTTACACACCTTACATACATAAGAAAATATCACATGAATCTCATATAAT 88399

QY 64 GGGTTAGTACTGTTTATATAAGAGAGTATAAAGATACATCATCATCTTTTGAGGCAATAA 123
DB 88398 ATGTAAATATTTGTGTATCAATTTTAAAAAACCCAAAAAATTTTCTTATGATCATGTAA 88339

QY 124 GCGAGGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTTAAAGTG 183
DB 88338 ATAATACATACCCATCGTTTAAAGTTTGGAAATTCACAGAAGACATGATGAAAAAATTA 88279

QY 184 ACCCCCTCCTTGA 197
DB 88278 AAATACCCATTAA 88265

RESULT 13
US-09-949-016-15524/c
; Sequence 15524, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15524
; LENGTH: 256176
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(256176)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15524/c
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	10.5%	Score 35.2	DB 3	Length 663
	Best Local Similarity	58.7%	Pred. No. 0.78		
	Matches	61	Conservative 0	Mismatches 43	Indels 0
				Gaps 0	
Qy	1	ATTATTAATTTTCATTAGATGCGGGTTTTTTTACTACAACTCAATATAGATGAACAGATG	60		
Db	567	ATTATTAACCTTTATTAGTTAAACCATTTATTAAATTGATCAATAATAATATAAAGCAATAA	508		
Qy	61	AATGGGTTAGTGACTGTTTATAAGAAGAGTAAATAAGATCTA	104		
Db	507	CATTAAATGATATAATAGTTATTATPAGAACCAAATGGAAGATCTA	464		

Search completed: July 2, 2005, 20:45:04
Job time : 68.0909 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 17:53:02 ; Search time 225.814 Seconds
(without alignments)
9332.708 Million cell updates/sec

```

Title: US-10-099-663-2
Perfect score: 336
Sequence: 1 attattatttttcattagata.....aqacagaacatggcattta 336

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 631374 seqs, 3136092125 residues
Total number of hits satisfying chosen parameters: 12626748

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```

Database : Published Applications NA:*
1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq:*
9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq:*

```

SUMMARIES

Result No.	Score	Query #		DB	ID	Description
		Match	Length			
1	336	100.0	336	16	US-10-099-663-2	Sequence 2, Appli
2	336	100.0	2381	16	US-10-099-663-1	Sequence 1, Appli
3	40.4	12.0	115327	20	US-10-719-993-6867	Sequence 6867, Ap
C 4	37.2	11.1	123526	21	US-09-510-186-11	Sequence 11, Appl
C 5	36.8	11.0	98439	21	US-10-741-600-17724	Sequence 17724, A
6	36.6	10.9	403	9	US-09-815-343-915	Sequence 915, App
7	36.6	10.9	403	18	US-10-097-105-915	Sequence 915, App

ALIGNMENTS

```

RESULT 1
US-10-099-663-2
; Sequence 2, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian GUT-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
US-10-099-663-2

```

8	35.2	10.8	8648	17	US-10-221-613-217	Sequence 217, App
9	35.8	10.7	714	21	US-10-706-635-64	Sequence 64, Appl
10	35.8	10.7	50000	21	US-10-706-635-26	Sequence 26, Appl
11	35.6	10.6	165097	22	US-10-737-082-77	Sequence 77, Appl
12	35.6	10.6	165097	22	US-10-765-790-77	Sequence 77, Appl
13	35.4	10.5	3673778	16	US-10-312-841-1	Sequence 1, Appl1
14	35.2	10.5	479	17	US-10-242-535A-47091	Sequence 47091, A
15	35.2	10.5	479	18	US-10-085-783A-47091	Sequence 47091, A
16	35.2	10.5	8622	15	US-10-311-455-2115	Sequence 2115, Ap
17	35.2	10.5	15714	15	US-10-311-455-1146	Sequence 1146, Ap
18	35.2	10.5	15714	19	US-10-433-793-88	Sequence 88, Appl
19	35	10.4	134738	13	US-10-087-192-502	Sequence 502, App
20	34.8	10.4	75899	9	US-09-854-883-243	Sequence 243, App
21	34.8	10.4	75899	17	US-10-360-510-243	Sequence 243, App
22	34.8	10.4	75899	23	US-11-008-747-243	Sequence 243, App
23	34.6	10.3	6072	15	US-10-311-455-4	Sequence 4, Appl1
24	34.4	10.2	1173	17	US-10-369-493-24780	Sequence 24780, A
25	34.4	10.2	10957	15	US-10-311-455-1084	Sequence 1084, Ap
26	34.2	10.2	598	19	US-10-767-701-24993	Sequence 24993, A
27	34.2	10.2	8605	15	US-10-311-455-508	Sequence 508, App
28	34.2	10.2	98844	21	US-10-467-182-10	Sequence 10, Appl
29	34.2	10.2	151870	19	US-10-741-601-5614	Sequence 5614, Ap
30	34.2	10.2	151870	21	US-10-741-600-17561	Sequence 17561, A
31	34	10.1	290547	19	US-10-367-094-77	Sequence 77, Appl
32	33.8	10.1	1028	19	US-10-767-795-4283	Sequence 4283, Ap
33	33.8	10.1	2542	13	US-10-027-632-111567	Sequence 111567, Sequence 111567,
34	33.8	10.1	2542	17	US-10-027-632-111567	Sequence 111567,
35	33.8	10.1	41227	17	US-10-374-780A-2169	Sequence 2169, Ap
36	33.8	10.1	127238	13	US-10-087-192-787	Sequence 787, App
37	33.6	10.0	449	18	US-10-434-599-51020	Sequence 51020, A
38	33.6	10.0	1210	13	US-10-037-632-216618	Sequence 216618, Sequence 216618,
39	33.6	10.0	1210	17	US-10-037-632-216618	Sequence 216618,
40	33.6	10.0	1770	13	US-10-027-632-259643	Sequence 259643, Sequence 259643,
41	33.6	10.0	1770	17	US-10-027-632-259643	Sequence 259643,
42	33.6	10.0	2119	21	US-10-764-420-2183	Sequence 2183, Ap
43	33.6	10.0	17527	15	US-10-311-455-1406	Sequence 1406, Ap
44	33.6	10.0	17527	18	US-10-240-454-28	Sequence 28, Appl
45	33.4	9.9	528	19	US-10-021-323-8131	Sequence 8131, Ap

QY 61 AATGGGTAGTGTCTTTATAAGAGAGTAAATAAGATACATCATCTTTGAGGCAA 120
Db 61 AATGGGTAGTGTCTTTATAAGAGAGTAAATAAGATACATCATCTTTGAGGCAA 120
QY 121 TAAGGGAGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTAAA 180
Db 121 TAAGGGAGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTAAA 180
QY 181 GTGACCCCTCTCTTGCACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
Db 181 GTGACCCCTCTCTTGCACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
QY 241 TGTAAATTCCTTCTGATAGGCTGTTCATAAATTCCTTTGCAAGCTCTGCTACTTA 300
Db 241 TGTAAATTCCTTCTGATAGGCTGTTCATAAATTCCTTTGCAAGCTCTGCTACTTA 300
QY 301 CCAGAAGTCTGCTTACAGACAGAAAGATGCGCATTTA 336
Db 301 CCAGAAGTCTGCTTACAGACAGAAAGATGCGCATTTA 336

RESULT 2

US-10-099-663-1
; Sequence 1, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Gut-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(1626)
; NAME/KEY: exon
; LOCATION: (1627)..(1693)
; NAME/KEY: Intron
; LOCATION: (1694)..(2322)
; NAME/KEY: exon
; LOCATION: (2333)..(2381)
US-10-099-663-1

Query Match 100.0%; Score 336; DB 16; Length 2381;
Best Local Similarity 100.0%; Pred. No. 3.1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTATTATTTTCATTAGATAGCCGGTTTTTTTACTACAACCTCAATAAGATGAACAGAATG 60
Db 1301 ATTATTATTTTCATTAGATAGCCGGTTTTTTTACTACAACCTCAATAAGATGAACAGAATG 1360
QY 61 AATGGGTAGTGTCTTTATAAGAGAGATAAATAAGATACATCATCTTTGAGGCAA 120
Db 1361 AATGGGTAGTGTCTTTATAAGAGAGATAAATAAGATACATCATCTTTGAGGCAA 1420
QY 121 TAAGGGAGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTAAA 180
Db 1421 TAAGGGAGGAGAGATTTCAGCAACAGTGTCTTACAAGTGGAAAAACAAGTTAAACTAAA 1480
QY 181 GTGACCCCTCTCTTGCACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 240
Db 1481 GTGACCCCTCTCTTGCACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCA 1540
QY 241 TGTAAATTCCTTCTGATAGGCTGTTCATAAATTCCTTTGCAAGCTCTGCTACTTA 300
Db 1541 TGTAAATTCCTTCTGATAGGCTGTTCATAAATTCCTTTGCAAGCTCTGCTACTTA 1600
QY 301 CCAGAAGTCTGCTTACAGACAGAAAGATGCGCATTTA 336

Db 1601 CCAGAAGTCTGCTTACAGACAGAAAGATGCGCATTTA 1636

RESULT 3

US-10-719-993-6867
; Sequence 6867, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6867
; LENGTH: 116327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(116327)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6867

Query Match 12.0%; Score 40.4; DB 20; Length 116327;
Best Local Similarity 55.8%; Pred. No. 4.8;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 197 ACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATCTATAAATGCTTTTCT 256
Db 70844 AAAAGATGATGCCCTCAACTAGTTTGACACAGTAGTTTAGCATGCATGAAGAAATACCT 70903
QY 257 GATAAGCCTGTTTCATAAATTCCTTTGCAAGCTCTGCTACTTACCAGAGTCTGCCTAC 316
Db 70904 GGTAGCTTTTTTAAAAATTCCTTCCCAAGTCTTACTTCTAAAAAAGAGTCAGTTTTT 70963
QY 317 AGACAGAAAGATGGCATT 334
Db 70964 AGAAGAAAAAGAGTATT 70981

RESULT 4

US-09-910-185-11/c
; Sequence 11, Application US/09910185
; Publication No. US20030083279A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLIOMA-ASSOCIATED ONCOGENE-3 EXPRESSION
; FILE REFERENCE: RTS-0258
; CURRENT APPLICATION NUMBER: US/09/910,185
; CURRENT FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 11
; LENGTH: 123526
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-910-185-11

Query Match 11.1%; Score 37.2; DB 10; Length 123526;
Best Local Similarity 53.4%; Pred. No. 40;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 12 CATTAGATAGCCGGTTTTTTTACTACAACCTCAATAAGATGAACAGATCAATGGGTAGT 71
Db 107264 CATTTCCCAATAGCCCTTTTAAATTAATGCTACAGAAAATGCACAGAAAGAGGGGTCT 107205
QY 72 GACTGTTTAAAGAGAGATTAATAAGATACATCATCTATCTATCTTTGAGGCAATAAAGGGAGGGA 131
Db 107204 GTGTACCCATTAAATAATAGATCAAGAAAAGAAAGAGAAAGCCAGAGAGGAGGGA 107145

Qy 132 GAGATTGAGCAAAACAGTGTGCTTACA 157
||| | | | | | | |
Db 107144 GGGAGAAAAACCGAAAGCATCCATGCA 107119

```

RESULT 5
US-10-741-600-17724/c
; Sequence 17724, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17724
; LENGTH: 98439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17724

```

Query Match	11.0%;	Score 36.8;	DB 21;	Length 98439;
Best Local Similarity	49.5%;	Pred. No. 47;		
Matches	95;	Conservative 0;	Mismatches 97;	Indels 0; Gaps 0;
QY	9	TTTCATTAGATAGCGGTTTTTTTTTACTACAACTCAAAATGAGATGAACAGATGAATGGGTT	68	
Db	49015	TCTCCTGGCTAACTCTTGCCACAATTTGAGCATCAAAATAAATAATAAAGGAAGGATT	48956	
QY	69	AGTGACTGTTTTAAAGAAAGAGTAATAAAGATACTATCATCATTTTGAGGCAATRAAGGAG	128	
Db	48955	ACAGCCCTGTTTTTAAAAAATAAGAAATCACAATCTACATTGATATAAGTAAATAATAACA	48996	
QY	129	GGAGAGATTCCAGCAACAGCTGTGCTTTACAGTGGAAACACAGTTTAACTTAAGTGACCCC	188	
Db	48895	TGGGGGAGAGGGGGAATTCCTTCCTTACAGAACTCAATGAATAAATACAAAAGGAATAA	48836	
QY	189	CCTCCTTGACAA	200	
Db	48835	TATTCCTTAGAAA	48824	

```

RESULT 6
US-09-815-343-915
; Sequence 915, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 915
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(403)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-915

```

Query Match 10.9%; Score 36.6; DB 9; Length 403;
Best Local Similarity 50.6%; Pred. No. 4.4;

	Matches	84;	Conservative	0;	Mismatches	82;	Indels	0;	Gaps	0;
Qy	26	TTTTTTACTACA	ACTCAA	ATAAGATGAAC	AGAAATGAATGGGTTAGTGACTCTTTTATAAAG	85				
Db	178	TGTTTAAAGTTAA	AGNAAAAGAGCTG	CAGAGTATTTATAAACTGCTCTTTAGAAA	237					
Qy	86	AAGAGTAATAA	GATACTATCATCT	TATTTGAGGCAATAAGGGAGGAGAGATTCAGCAAC	145					
Db	238	AACAAGCAAGA	GACCATTTTGACCATATGAATG	GAAGGGAAGAAAGTATTATAGAAAC	297					
Qy	146	AGTGTGCTTACA	AGTGGAAAAACAAGTTAAACTAAAGTGA	CCCCCCT	191					
Db	298	TTTGCTAGTTNA	AAAAAATAAAAAAATACTTNGCNGAAC	CCCCCT	343					

```

RESULT 7
US-10-097-105-915
; Sequence 915, Application US/10097105
; Publication No. US20040037842A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 915
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 308..330, 334, 356, 359, 368, 369, 379, 383, 387,
; OTHER INFORMATION: n = A,T,C or G
US-10-097-105-915

```

	Query Match	10.9%;	Score 36.6;	DB 18;	Length 403;
	Best Local Similarity	50.6%;	Pred. No. 4.4;		
	Matches	84;	Conservative 0;	Mismatches 82;	Indels 0; Gaps 0;
QY	26	TTTTTTTACTACAAC	TCAAAATAGATGAA	CAGATGAATGGGTTAGTGACTGTTTATATAAG	85
DB	178	TTGTTTTAAAGTTAAAGAAAAA	GAGCTGCAGAGTATTTAT	AAAAACTGCTCTTTTAGAAAAA	237
QY	86	AAGAGTAATAAAGATACTATCAT	TTTCAGGCAATAAGGAGGGAGAGATT	CAGCAAC	145
DB	238	AACAAGCAGAGACCACTTTGAC	CCATATGATGGAAGGGAAGAAAGTATTATAGAAAC		297
QY	146	AGTGTGCTTTACAAGTGGAAAA	CAAGTTAAACTAAAGTGACCCCCCT		191
DB	298	TTTGCTAGTTTAAAAA	AAAAAATAAAAAAACTTNGCCNGAACCCCT		343

```

RESULT 8
US-10-221-613-217
; Sequence 217, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 13:41:36 ; Search time 1215.63 Seconds
(without alignments)
10520.918 Million cell updates/sec

Title: US-10-099-663-2

Perfect score: 336

Sequence: 1 attattattttcattagata.....agacagaagatggcattta 336

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0.

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.8	12.4	500	1	AU088319
2	41.4	12.3	444	5	BW139785
3	41.4	12.3	504	5	BW269615
4	41.4	12.3	574	1	AV680413
5	41.4	12.3	720	5	BW135968
6	41.4	12.3	725	5	BW186283
7	41.4	12.3	730	5	BW140184
8	41.4	12.3	732	5	BW127139
9	41.4	12.3	759	1	AV877102
10	41.4	12.3	767	5	BW082303
11	41.4	12.3	774	5	BW141206
12	41	12.2	598	7	CF250818
13	40.6	12.1	594	8	AZ305625
14	40.6	12.1	732	9	AG596239
15	40.6	12.1	781	9	AG470649
16	40.4	12.0	554	5	BW009156
17	40.2	12.0	522	8	AZ431375
18	40	11.9	879	8	BH137539
19	40	11.9	943	8	BH147340
20	39.8	11.8	472	5	BW093545
21	39.8	11.8	507	5	BW234156
22	39.8	11.8	688	5	BW133492
23	39.8	11.8	760	5	BW399296
24	39.2	11.7	1101	9	CNS00EQ6

CE141435 tigr-g88-
BW403626 BW403626
BH135160 ENT0B35TF
BW124142 BW124142
BW254209 BW254209
BZ393910 EINCW68TR
BH146427 ENT0H92TR
AL104456 Drosophil
BH784350 BH784350
BH973910 oad52e01.
BZ011773 oeh48403.
CK141433 Reverse B
CK383193 la117908.
AZ002591 RPCI-23-2
BF286550 EST451141
AZ549279 ENTEG46TF
AL107995 Drosophil
AL106572 Drosophil
CL81587 abf51c07.
BZ948176 CH240.113
CD825484 BN25.0600

ALIGNMENTS

RESULT 1
AU088319 500 bp mRNA linear EST 27-JAN-2001
LOCUS AU088319 Sugano Malaria cDNA library Plasmodium falciparum 3D7 cDNA
DEFINITION Clone XPfn6924, mRNA sequence.
ACCESSION AU088319
VERSION AU088319.1 GI:12390460
KEYWORDS EST.
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 500)
Watanabe,J., Sasaki,M., Suzuki,Y. and Sugano,S.
FULL-malaria: a database for a full-length enriched cDNA library
from human malaria parasite, Plasmodium falciparum
Nucleic Acids Res. 29 (1), 70-71 (2001)
JOURNAL 20574754
MEDLINE 11125052
PUBMED
COMMENT Contact: Junichi Watanabe
Institute of Medical Science
The University of Tokyo, Department of Parasitology
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Tel: 81-3-5449-5378
Fax: 81-3-5449-5410
Email: jwatanabe@med.s.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
Sugano,S. Construction and characterization of a full
length-enriched and a 5'-end-enriched cDNA library Gene 200 (1-2),
149-156 (1997).

FEATURES

source
1..500
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"
/isolate="3D7"
/db_xref="taxon:36329"
/clone="XPfn6924"
/dev_stage="erythrocytic stage"
/clone_lib="Sugano Malaria cDNA library"

ORIGIN

Query Match 12.4%; Score 41.8; DB 1; Length 500;
Best Local Similarity 50.2%; Pred. No. 0.93;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
14 TTAGATAGCCGGTGTCTTACTACAACTCAATGAATGAATGAATGGTTAGTGA 73
232 TTAATTATGATGATATATAGAAATGGAATATAAATAAATAACGGAATCATTAA 291

QY 74 CTGTTTATAAGAGAGTAATAAGATACATCATCTTTGAGGCAATTAAGGAGGAGA 133
 Db 292 ATGTCTTTTATAAGAAAAATCAAATGAGGAAGATTCTTTACAAAAAAGTTGAATGA 351
 QY 134 GATTCAGCAACAGTGTCTTACAGTGGAAACAAAGTTAACTAAAGTGACCCCTCC 193
 Db 352 CAAAGTAATTAACGTGCMAACAAAGGGAACAAAAAAGAAATGTTATCACAAAGG 411
 QY 194 TTGACAAGATCAATGCCAGTTGA 218
 Db 412 TACATAAAACAAACAAAATTTGA 436

RESULT 2
 BW139785
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BW139785 444 bp mRNA linear EST 03-NOV-2002
 BW139785 Nori Satoh unpublished cDNA library, gastrula and neurula
 Ciona intestinalis cDNA clone rcign053020 3', mRNA sequence.
 BW139785
 EST.
 GI:24496807
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 444)
 Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satohe@acidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1. .444
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign053020"
 /tissue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 444;
 Best Local Similarity 46.4%; Pred. No. 1.2;
 Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATTAAGATGAACAGATGAATGGGTAGTACTGTTTATAAAGAGAGTAAT 94
 Db 17 ATAACTGTATTACGAACAGACACATATGTGTTCCAAACGGTCAACAATTAAGTAAT 76
 QY 95 AAAGATATCATCATCTTTGAGCAATAAGGGAGGAGAGATTACGACAAACAGTGTGCTT 154
 Db 77 GACAGTTCCAATACCTTTCTTGTAAAAAATATGACCCCTGCTTCATTTTACATTGCACCA 136
 QY 155 ACAAGTGGAAAAACAGTTAACTAAAGTGACCCCTCTTGCACAAGATCAATGCCACAG 214
 Db 137 CGAAATGGAAATAGGTTTAAATAGTAAACATTTAAATAATGTTACCGAAATGCACCCCTG 196
 QY 215 TTGAGCTTTAGCCAGCCACATCATCATGTAAATGCTTTCCTTGATAAGCTGTTTCATAAA 274
 Db 197 CTAAGCATCTGCTATTGCTGTATTAATTCCTGTGTGTTCTTGCACAACATAGTTGCAAC 256
 QY 275 TTCTCTTTGCAAGCTCTGCTACTTACCAAGAGTTCGCTACAGACAGAAA 325
 Db 257 AGCTTTCTGTAAATCTCCACCACCTTGCAATTCAGTTCTGCTTCTGCTCTTAA 307

RESULT 3
 BW269615/c

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BW269615 504 bp mRNA linear EST 09-NOV-2002
 BW269615 Nori Satoh unpublished cDNA library, gastrula and neurula
 Ciona intestinalis cDNA clone cign053020 5', mRNA sequence.
 BW269615
 EST.
 GI:24849533
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 504)
 Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satohe@acidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1. .504
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cign053020"
 /tissue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 504;
 Best Local Similarity 46.4%; Pred. No. 1.2;
 Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATTAAGATGAACAGATGAATGGGTAGTACTGTTTATAAAGAGAGTAAT 94
 Db 414 ATAACTGTATTACGAACAGACACATATGTGTTCCAAACGGTCAACAATTAAGTAAT 355
 QY 95 AAAGATATCATCATCTTTGAGCAATAAGGGAGGAGAGATTACGACAAACAGTGTGCTT 154
 Db 354 GACAGTTCCAATACCTTTCTTGTAAAAAATATGACCCCTGCTTCATTTTACATTGCACCA 295
 QY 155 ACAAGTGGAAAAACAGTTAACTAAAGTGACCCCTCTTGCACAAGATCAATGCCACAG 214
 Db 294 CGAAATGGAAATAGGTTTAAATAGTAAACATTTAAATAATGTTACCGAAATGCACCCCTG 235
 QY 215 TTGAGCTTTAGCCAGCCACATCATCTGTAAATGCTTTCCTTGATAAGCTGTTTCATAAA 274
 Db 234 CTAAGCATCTGCTATTGCTGTATTAATTCCTGTGTGTTCTTGCACAACATAGTTGCAAC 175
 QY 275 TTCTCTTTGCAAGCTCTGCTACTTACCAAGAGTTCGCTACAGACAGAAA 325
 Db 174 AGCTTTCTGTAAATCTCCACCACCTTGCAATTCAGTTCTGCTTCTGCTCTTAA 124

RESULT 4
 AV680413

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AV680413 574 bp mRNA linear EST 05-OCT-2000
 AV680413 Nori Satoh unpublished cDNA library Ciona intestinalis
 cDNA clone rcitb12j8 3', mRNA sequence.
 AV680413
 EST.
 GI:10118412
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 574)
Sato, N., Satou, Y., Kohara, Y. and Shin-i, T.
Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1. .574
Location/Qualifiers

/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcitb1218"
/tissue_type="whole animal"
/dev_stage="tailbud"
/clone_lib="Nori Satoh unpublished cDNA library"

ORIGIN

Query Match 12.3%; Score 41.4; DB 1; Length 574;
Best Local Similarity 46.4%; Pred. No. 1.2; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 156;

QY 35 ACAACTCAAAATAGATGAACAGAGATGAATGGGTAGTACTGTTTATAAAGAGAGTAAT 94
DB 36 ATAACTGTATTACGAAAGCAGACATATGTGTTCCAAACGGTCAACAATAGGTAAAT 95
QY 95 AAGATATCTATCATCTTTGAGGCAATTAAGGAGGAGAGATTACGAAACAGTGTGCTT 154
DB 96 GACAGTTCCTCAATACCTTTCTGTTAAATAATATGACCCCTGCTTCATTTACATTGCACCA 155
QY 155 ACAAGTGGAAACAGTTAACTAAAGTGACCCCTCTTGACAGATCAATGCCACAG 214
DB 156 CGAAATGGAATAGTGTAAATAGTAAACATTAATAATGTTACCGAAATGCCCCCTG 215
QY 215 TTGAGCTTTAGCCAGCCACATCATGTAAATGCTTTCTGATAGAGCTGTTTCATAAA 274
DB 216 CTAAAGCATCTGCTATTGCTGTTAAATTTCTGTGTTCTTGACACATAGTTCAC 275
QY 275 TTCTTTTGAAGAGTCTGCTACTTACCAAGAGTCTGCCCTACAGACAGAAA 325
DB 276 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTCTGCTCTAA 326

RESULT 5

BW135968
LOCUS
DEFINITION
BW135968 Nori Satoh unpublished cDNA library, linear EST 02-NOV-2002
Ciona intestinalis cDNA clone rcign042b07 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BW135968.1 GI:24492367
EST.
Ciona intestinalis

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sato, Y., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1. .720
Location/Qualifiers

/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcit035k14"
/tissue_type="heart"

QY 35 ACAACTCAAAATAGATGAACAGAGATGAATGGGTAGTACTGTTTATAAAGAGTAAT 94

DB 36 ATAACTGTATTACGAAAGCAGACATATGTGTTCCAAACGGTCAACAATAGGTAAAT 95

QY 95 AAGATATCTATCATCTTTGAGGCAATTAAGGAGGAGAGATTACGAAACAGTGTGCTT 154

DB 96 GACAGTTCCTCAATACCTTTCTGTTAAATAATATGACCCCTGCTTCATTTACATTGCACCA 155

QY 155 ACAAGTGGAAACAGTTAACTAAAGTGACCCCTCTTGACAGATCAATGCCACAG 214

DB 156 CGAAATGGAATAGTGTAAATAGTAAACATTAATAATGTTACCGAAATGCCCCCTG 215

QY 215 TTGAGCTTTAGCCAGCCACATCATGTAAATGCTTTCTGATAGAGCTGTTTCATAAA 274

DB 216 CTAAAGCATCTGCTATTGCTGTTAAATTTCTGTGTTCTTGACACATAGTTCAC 275

QY 275 TTCTTTTGAAGAGTCTGCTACTTACCAAGAGTCTGCCCTACAGACAGAAA 325

DB 276 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTCTGCTCTAA 326

ORIGIN
Query Match 12.3%; Score 41.4; DB 5; Length 720;
Best Local Similarity 46.4%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 156;

QY 35 ACAACTCAAAATAGATGAACAGAGATGAATGGGTAGTACTGTTTATAAAGAGAGTAAT 94
DB 24 ATAACTGTATTACGAAAGCAGACATATGTGTTCCAAACGGTCAACAATAGGTAAAT 83
QY 95 AAGATATCTATCATCTTTGAGGCAATTAAGGAGGAGAGATTACGAAACAGTGTGCTT 154
DB 84 GACAGTTCCTCAATACCTTTCTGTTAAATAATATGACCCCTGCTTCATTTACATTGCACCA 143
QY 155 ACAAGTGGAAACAGTTAACTAAAGTGACCCCTCTTGACAGATCAATGCCACAG 214
DB 144 CGAAATGGAATAGTGTAAATAGTAAACATTAATAATGTTACCGAAATGCCCCCTG 203
QY 215 TTGAGCTTTAGCCAGCCACATCATGTAAATGCTTTCTGATAGAGCTGTTTCATAAA 274
DB 204 CTAAAGCATCTGCTATTGCTGTTAAATTTCTGTGTTCTTGACACATAGTTCAC 263
QY 275 TTCTTTTGAAGAGTCTGCTACTTACCAAGAGTCTGCCCTACAGACAGAAA 325
DB 264 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTCTGCTCTAA 314

RESULT 6

BW186283/c
LOCUS
DEFINITION
BW186283 Nori Satoh unpublished cDNA library, linear EST 05-NOV-2002
Ciona intestinalis cDNA clone rcit035k14 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BW186283.1 GI:24576615
EST.
Ciona intestinalis

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sato, Y., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1. .725
Location/Qualifiers

/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcit035k14"
/tissue_type="heart"

ORIGIN
Query Match 12.3%; Score 41.4; DB 5; Length 725;
Best Local Similarity 46.4%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 156;

QY 35 ACAACTCAAAATAGATGAACAGAGATGAATGGGTAGTACTGTTTATAAAGAGTAAT 94
DB 36 ATAACTGTATTACGAAAGCAGACATATGTGTTCCAAACGGTCAACAATAGGTAAAT 83
QY 95 AAGATATCTATCATCTTTGAGGCAATTAAGGAGGAGAGATTACGAAACAGTGTGCTT 154
DB 84 GACAGTTCCTCAATACCTTTCTGTTAAATAATATGACCCCTGCTTCATTTACATTGCACCA 143
QY 155 ACAAGTGGAAACAGTTAACTAAAGTGACCCCTCTTGACAGATCAATGCCACAG 214
DB 144 CGAAATGGAATAGTGTAAATAGTAAACATTAATAATGTTACCGAAATGCCCCCTG 203
QY 215 TTGAGCTTTAGCCAGCCACATCATGTAAATGCTTTCTGATAGAGCTGTTTCATAAA 274
DB 204 CTAAAGCATCTGCTATTGCTGTTAAATTTCTGTGTTCTTGACACATAGTTCAC 263
QY 275 TTCTTTTGAAGAGTCTGCTACTTACCAAGAGTCTGCCCTACAGACAGAAA 325
DB 264 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTCTGCTCTAA 314

Db 304 ATAACTGTATTACGAAAGCAGAACATATGTGTTCACAAAGGTCAACAATAAGGTAAT 245

QY 95 AAGATACTATCATCTTTGAGCAATAAGGGAGGAGAGATTACGAAACAGTGTGCTT 154

Db 244 GACAGTTCCAATACCTTTTCTGTAAATAATATGACCTGCTTCATTTTACATGTCACCA 185

QY 155 ACAAGTGGAAACAAGTTAACTAAAGTGACCCCTCTTGACAAGATCAATGCCACAG 214

Db 184 CGAAATGGAAATAGGTGTAAATAATAGTAAACATTAATAATTTGTACCGAAATGCCCCCTG 125

QY 215 TTGAGCTTTAGCCAGGCACATCATCATGTAAATGCTTCCCTGATAAGCTGTTCATAAA 274

Db 124 CTAAGCATCTGCTATTGTGTATTTAATCTGTGTCTTCTTGACAACATAGTTGCAAC 65

QY 275 TTCTCTTTGAAAGCTCTCTACTTACCAAGATCTGCTTACAGACAGAA 325

Db 64 AGCTTTCTGTAAATCTCCACCACTTGCAATTCAGTTCTGCTTCTGCTCTTAA 14

RESULT 7

BW140184 730 bp mRNA linear EST 03-NOV-2002

LOCUS BW140184 Nori Satoh unpublished cDNA library, gastrula and neurula

DEFINITION Ciona intestinalis cDNA clone rcign055e20 3', mRNA sequence.

ACCESSION BW140184

VERSION 1

KEYWORDS EST.

SOURCE GI:24497206

ORGANISM Ciona intestinalis

Phlebobranchia, Clonidae; Ciona.

REFERENCE

AUTHORS Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.

TITLE Expressed genes in Ciona intestinalis (2002c)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

1. .730

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="rcign055e20"

/tissue_type="whole body"

/dev_stage="gastrula and neurula"

/clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 730;

Best Local Similarity 46.4%; Pred. No. 1.3; Mismatches 0; Gaps 0;

Matches 135; Conservative 0; Indels 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATAAGATGAACAGAAATGAATGGGTAGTGAAGTGTATTAAGAAGAGTAAT 94

Db 40 ATAACTGTATTACGAAAGCAGAACATATGTGTTCACAAAGGTCAACAATAAGGTAAT 99

QY 95 AAGATACTATCATCTTTGAGCAATAAGGGAGGAGAGATTACGAAACAGTGTGCTT 154

Db 100 GACAGTTCCAATACCTTTTCTGTAAATAATATGACCTGCTTCATTTTACATGTCACCA 159

QY 155 ACAAGTGGAAACAAGTTAACTAAAGTGACCCCTCTTGACAAGATCAATGCCACAG 214

Db 160 CGAAATGGAAATAGGTGTAAATAATAGTAAACATTAATAATTTGTACCGAAATGCCCCCTG 219

QY 215 TTGAGCTTTAGCCAGGCACATCATCATGTAAATGCTTCCCTGATAAGCTGTTCATAAA 274

Db 220 CTAAGCATCTGCTATTGTGTATTTAATCTGTGTCTTCTTGACAACATAGTTGCAAC 279

Db 304 ATAACTGTATTACGAAAGCAGAACATATGTGTTCACAAAGGTCAACAATAAGGTAAT 245

QY 95 AAGATACTATCATCTTTGAGCAATAAGGGAGGAGAGATTACGAAACAGTGTGCTT 154

Db 244 GACAGTTCCAATACCTTTTCTGTAAATAATATGACCTGCTTCATTTTACATGTCACCA 185

QY 155 ACAAGTGGAAACAAGTTAACTAAAGTGACCCCTCTTGACAAGATCAATGCCACAG 214

Db 184 CGAAATGGAAATAGGTGTAAATAATAGTAAACATTAATAATTTGTACCGAAATGCCCCCTG 125

QY 215 TTGAGCTTTAGCCAGGCACATCATCATGTAAATGCTTCCCTGATAAGCTGTTCATAAA 274

Db 124 CTAAGCATCTGCTATTGTGTATTTAATCTGTGTCTTCTTGACAACATAGTTGCAAC 65

QY 275 TTCTCTTTGAAAGCTCTCTACTTACCAAGATCTGCTTACAGACAGAA 325

Db 64 AGCTTTCTGTAAATCTCCACCACTTGCAATTCAGTTCTGCTTCTGCTCTTAA 14

RESULT 8

BW127139 732 bp mRNA linear EST 02-NOV-2002

LOCUS BW127139 Nori Satoh unpublished cDNA library, gastrula and neurula

DEFINITION Ciona intestinalis cDNA clone rcign013j16 3', mRNA sequence.

ACCESSION BW127139

VERSION 1

KEYWORDS EST.

SOURCE GI:24483538

ORGANISM Ciona intestinalis

Phlebobranchia, Clonidae; Ciona.

REFERENCE

AUTHORS Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.

TITLE Expressed genes in Ciona intestinalis (2002c)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

1. .732

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="rcign013j16"

/tissue_type="whole body"

/dev_stage="gastrula and neurula"

/clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 732;

Best Local Similarity 46.4%; Pred. No. 1.3; Mismatches 0; Gaps 0;

Matches 135; Conservative 0; Indels 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATAAGATGAACAGAAATGAATGGGTAGTGAAGTGTATTAAGAAGAGTAAT 94

Db 20 ATAACTGTATTACGAAAGCAGAACATATGTGTTCACAAAGGTCAACAATAAGGTAAT 79

QY 95 AAGATACTATCATCTTTGAGGCAATAAGGGAGGAGAGATTACGAAACAGTGTGCTT 154

Db 80 GACAGTTCCAATACCTTTTCTGTAAATAATATGACCTGCTTCATTTTACATGTCACCA 139

QY 155 ACAAGTGGAAACAAGTTAACTAAAGTGACCCCTCTCTTGACAAGATCAATGCCACAG 214

Db 140 CGAAATGGAAATAGGTGTAAATAATAGTAAACATTAATAATTTGTACCGAAATGCCCCCTG 199

QY 215 TTGAGCTTTAGCCAGGCACATCATCATGTAAATGCTTTCCTGATAAGCTGTTCATAAA 274

Db 200 CTAAGCATCTGCTATTGTGTATTTAATCTGTGTCTTCTTGACAACATAGTTGCAAC 259

QY 275 TTCTCTTTGAAAGCTCTGCTACTTACCAAGATCTGCTTACAGACAGAA 325

Db 260 AGCTTTCTGTAAATCTCCACCACTTGCAATTCAGTTCTGCTTCTGCTCTTAA 310

RESULT 9

AV877102 759 bp mRNA linear EST 08-NOV-2001

LOCUS AV877102 Nori Satoh unpublished cDNA library, tailbud embryo Ciona

DEFINITION intestinalis cDNA clone rcitb31k05 3', mRNA sequence.

ACCESSION AV877102

VERSION AV877102.1 GI:16864626

KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE
1 (bases 1 to 759)
Sato, N., Satou, Y., Kohara, Y. and Shin-i, T.
Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..759
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcitb31k05"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud embryo"

ORIGIN
Query Match 12.3%; Score 41.4; DB 1; Length 759;
Best Local Similarity 46.4%; Pred. No. 1.3;
Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATAAGATGAACAGAAATGAATGGGTTAGTCTGTTTAAAGAGAGTAAT 94
|||||
Db 453 ATAACGTGTTACGAAAGCAGACATATGTGTCCAAACGGTCAACAATAAGGTAAT 512
|||||
QY 95 AAGAGATCTATCATCATCTATTTGAGGCAATAAGGGAGGAGAGATTTCAGCAACAGTGTGCTT 154
|||||
Db 513 GACAGTTCCAAATACCTTTCTGTGTAATAATATGACCCCTGCTTACATTCATTCGACCA 572
|||||
QY 155 ACAAGTGGAAACAGATTAACTAAAGTGACCCCTCTTCTGATAGAGCTGTTCATAA 214
|||||
Db 573 CGAAATGGAATAGTGTGTTAAATAGTAAACATTTAAACATTTGTTACCGAAATGCACCCCTG 632
|||||
QY 215 TTGAGCTTTAGCAGCCACATCATCATGTAATGCTTTCTCTGATAGCTGTTCATAA 274
|||||
Db 633 CTAAAGCATCTGCTATGCTGTAATTAATCTGTGTCTTCTGACACATAGTTCGAAC 692
|||||
QY 275 TTCTCTTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTTACAGACAGAAA 325
|||||
Db 693 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTTCTGCTCTTAA 743
|||||

RESULT 10
BW082303 767 bp mRNA linear EST 22-OCT-2002
LOCUS
DEFINITION
BW082303 Nori Satoh unpublished cDNA library, egg Ciona
intestinalis cDNA clone rcieg088c22 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE
1 (bases 1 to 767)
Sato, Y., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..767
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcieg088c22"
/tissue_type="whole animal"
/dev_stage="egg"
/clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Query Match 12.3%; Score 41.4; DB 5; Length 767;
Best Local Similarity 46.4%; Pred. No. 1.3;
Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATAAGATGAACAGAAATGAATGGGTTAGTCTGTTTAAAGAGAGTAAT 94
|||||
Db 27 ATAACGTGTTACGAAAGCAGACATATGTGTCCAAACGGTCAACAATAAGGTAAT 86
|||||
QY 95 AAGAGATCTATCATCATCTATTTGAGGCAATAAGGGAGGAGAGATTTCAGCAACAGTGTGCTT 154
|||||
Db 87 GACAGTTCCAAATACCTTTCTGTGTAATAATATGACCCCTGCTTACATTCATTCGACCA 146
|||||
QY 155 ACAAGTGGAAACAGATTAACTAAAGTGACCCCTCTTCTGATAGAGCTGTTCATAA 214
|||||
Db 147 CGAAATGGAATAGTGTGTTAAATAGTAAACATTTAAACATTTGTTACCGAAATGCACCCCTG 206
|||||
QY 215 TTGAGCTTTAGCAGCCACATCATCATGTAATGCTTTCTCTGATAGAGCTGTTCATAA 274
|||||
Db 207 CTAAAGCATCTGCTATGCTGTAATTAATCTGTGTCTTCTGACACATAGTTCGAAC 266
|||||
QY 275 TTCTCTTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTTACAGACAGAAA 325
|||||
Db 267 AGCTTTCTGTAATCTCCACCACTTGCAATTCAGTTCTGCTTCTGCTCTTAA 317
|||||

RESULT 11
BW141206 774 bp mRNA linear EST 03-NOV-2002
LOCUS
DEFINITION
BW141206 Nori Satoh unpublished cDNA library, gastrula and neurula
Ciona intestinalis cDNA clone rcign058h04 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE
1 (bases 1 to 774)
Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..774
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcign058h04"
/tissue_type="whole body"
/dev_stage="gastrula and neurula"
/clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

ORIGIN


```
Db      388 TAGGAGTGGGGGAAAAGAGAAAGTTTAAAAAAGAA 350
|||||
RESULT 14
AG596239/c
LOCUS   AG596239          732 bp    DNA    linear    GSS 05-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-526L09.TJ, genomic survey
sequence.
ACCESSION AG596239
VERSION   AG596239.1 GI:48357069
KEYWORDS  GSS.
SOURCE    Mus musculus molossinus
ORGANISM  Mus musculus molossinus
REFERENCE 1
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     BAC end Sequences of Library MSMg01
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 732)
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunlya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY      Vector      : pBACe3.6
R.Site 1     : EcoRI
R.Site 2     : EcoRI
FEATURES     Location/Qualifiers
source       1..732
             /organism="Mus musculus molossinus"
             /mol_type="genomic DNA"
             /sub_species="molossinus"
             /db_xref="taxon:57486"
             /clone="MSMg01-526L09.TJ"
             /sex="male"
             /tissue_type="mixture of kidney and spleen"
             /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      12.1%; Score 40.6; DB 9; Length 732;
Best Local Similarity 53.5%; Pred. No. 2.2;
Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY      22 CCGGTTTTTCTACAACTCAATAGATGAACAGATGAATGGGTTAGTACTGTTTAT 81
|||||
Db      656 CTGATTATTTCATCAAGGAAAGAAAACAGGGGAATAAGGAGAAAGTCAGAAAGAA 597
|||||
QY      82 AAAGAAGAGTAATAAGATATCTATCATCTTTGAGGCAATAAGGGAGGAGAGATTTCAGC 141
|||||
Db      596 AAAGAAGATGAAGAAAGAAAAGGAGTAAGGAGGAGGAGGACGCGGAGG 537
|||||
QY      142 AAACAGTGTCTTACAGTGGAAAACAAGTTAACTAAA 180
|||||
Db      536 TAGGAGTGGGGGAAAAGAAAGTAAAAAAGAA 498
|||||

RESULT 15
AG470649/c
LOCUS   AG470649          781 bp    DNA    linear    GSS 04-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-361M24.TJ, genomic survey
sequence.
ACCESSION AG470649
VERSION   AG470649.1 GI:48173781
KEYWORDS  GSS.
SOURCE    Mus musculus molossinus
ORGANISM  Mus musculus molossinus
REFERENCE 1
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     BAC end Sequences of Library MSMg01
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 781)
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunlya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY      Vector      : pBACe3.6
R.Site 1     : EcoRI
R.Site 2     : EcoRI
FEATURES     Location/Qualifiers
source       1..781
             /organism="Mus musculus molossinus"
             /mol_type="genomic DNA"
             /sub_species="molossinus"
             /db_xref="taxon:57486"
             /clone="MSMg01-361M24.TJ"
             /sex="male"
             /tissue_type="mixture of kidney and spleen"
             /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      12.1%; Score 40.6; DB 9; Length 781;
Best Local Similarity 53.5%; Pred. No. 2.2;
Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY      22 CCGGTTTTTCTACAACTCAATAGATGAACAGATGAATGGGTTAGTACTGTTTAT 81
|||||
Db      343 CTGATTATTTCATCAAGGAAAGAAAACAGGGGAATAAGGAGAAAGTCAGAAAGAA 284
|||||
QY      82 AAAGAAGAGTAATAAGATATCTATCATCTTTGAGGCAATAAGGGAGGAGAGATTTCAGC 141
|||||
Db      283 AAAGAAGATGAAGAAAGAAAAGGAGTAAGGAGGAGGAGGACGCGGAGG 224
|||||
QY      142 AAACAGTGTCTTACAGTGGAAAACAAGTTAACTAAA 180
|||||
Db      223 TAGGAGTGGGGGAAAAGAAAGTAAAAAAGAA 185
|||||

Search completed: July 2, 2005, 20:36:46
Job time : 1223.63 secs
```

THIS PAGE BLANK (USPTO)